

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS QUÍMICAS

Departamento de Bioquímica y Biología Molecular I



TESIS DOCTORAL

**RMN y reconocimiento molecular: relación estructura-
actividad de péptidos implicados en el dolor e interacciones
proteína-carbohidrato**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Luis Pablo Calle Jiménez

Directores

Jesús Jiménez Barbero
Francisco Javier Cañada Vicinay

Madrid, 2017



UNIVERSIDAD COMPLUTENSE DE MADRID

Facultad de Ciencias Químicas

Departamento de Bioquímica y Biología Molecular I



RMN y Reconocimiento Molecular: Relación estructura-actividad de péptidos implicados en el dolor e interacciones proteína-carbohidrato

Memoria para optar al grado de doctor presentada por

Luis Pablo Calle Jiménez

Madrid, 2015

Esta Tesis ha sido realizada en el Centro de Investigaciones Biológicas
(CIB-CSIC)

Codirectores

Dr. Jesús Jiménez Barbero & Dr. Francisco Javier Cañada Vicinay

A mi familia

Agradecimientos

Esta Tesis Doctoral ha sido realizada en el Centro de Investigaciones Biológicas (CIB-CSIC), bajo la dirección del Prof. Jesús Jiménez Barbero y del Prof. Francisco Javier Cañada Vicinay, a quienes quiero agradecer haberme dado la oportunidad de formar parte de su grupo de investigación. Durante mi Tesis me han facilitado todas las herramientas necesarias para poder mejorar mi formación académica con la asistencia a numerosos cursos y congresos, así como con colaboraciones con grupos punteros en distintos campos que han enriquecido mi experiencia profesional. Su excelencia profesional ha estado siempre acompañada de su calidad humana, reflejándose en un ambiente laboral que ha conseguido compaginar enriquecedoras discusiones científicas con vivencias extraordinarias.

También quiero agradecer a todos los compañeros que han formado parte del grupo *RMN y reconocimiento molecular* del CIB-CSIC en algún momento durante este periodo de mi formación: Ana, Anita, Álvaro, Ángeles, Blanca, Carmen, Dolores, Filipa, Iñaki, João, José Juan, Karla, Khouzaima, Lidia, Luca, María, Pilar, Silvia, Valentina y Virginia.

La colaboración con distintos grupos de investigación me ha permitido realizar un trabajo más completo que ha enriquecido mi formación con el conocimiento de diversas áreas científicas. Quisiera agradecer a todos los investigadores con los que he tenido la oportunidad de compartir el trabajo realizado en esta tesis doctoral: Gregorio Valencia, Gemma Arsequell y Mónica Rosa (IQAC-CSIC en Barcelona); Catherine Rougeot (Departamento de *Biochimie Structurale et Cellulaire* del Instituto Pasteur de Paris); Raquel Rodríguez, Verónica González y Katherine Barreto (Departamento de *Bioquímica y Biología Molecular* de la Facultad de Medicina de la Universidad de Salamanca); Jesús T. Vázquez, Carlos Mayato y Rosa L. Dorta (Instituto Universitario de Bio-Organica Antonio González de la Universidad de La Laguna); Niels-Christian Reichardt, Begoña Echevarría y Sonia Serna (CIC BiomaGUNE en San Sebastián); Antonio Franconetti (Departamento de *Química Orgánica* de la Universidad de Sevilla); y, Oscar Varela, María Laura Uhrig, Alejandro J. Cagnoni y Verónica E. Manzano (CIHIDECAR-CONICET de la Universidad de Buenos Aires).

Igualmente me gustaría agradecer a Marta Bruix y a Manolis Matzapetakis el acceso que me proporcionaron para realizar medidas de RMN en los espectrómetros de 800 MHz en el Instituto de Química Física Rocasolano (Madrid) y en el ITQB (Oeiras, Portugal), respectivamente.

El último paso de esta tesis no hubiera sido posible sin el apoyo de mis compañeros de trabajo en Lilly (Alcobendas). En concreto, me gustaría agradecer su apoyo y ánimo diario a Aida, Francisco, Juan Félix, Manuel y Paloma.

Algunas personas que he conocido durante esta etapa profesional, han pasado a formar parte de mi vida personal. No quiero olvidarme de agradecerles su aliento durante la escritura de esta tesis. Además, este último empujón no hubiera sido posible sin la amistad de aquellos que han alegrado las desconexiones durante la escritura de la tesis. Me gustaría agradecer este apoyo a Beatriz, Esther, Diana, Fermín, Inés, Irene, Javi, Laura, Marco, María, Nerea, Pablo, Patxi, Sara y Teresa.

Por último, me gustaría agradecer y dedicar esta tesis al apoyo incondicional de mi familia. No quisiera olvidar la cálida bienvenida de mi familia política, el ejemplo de mi abuela Juana, la confianza generosa de mis padres y la complicidad de mi hermano. Pero el principal apoyo y la comprensión han sido proporcionados por mi pareja. Mi marido ha aguantado estoicamente la presión y la soledad que padece un compañero en momentos como este.

Índice

Agradecimientos	5
Índice	7
Abreviaturas.....	11
Resumen	13
Abstract.....	17
1. RMN y Reconocimiento Molecular	21
1.1. Péptidos y Carbohidratos.....	21
1.1.1. Péptidos.....	21
1.1.1.1. Enlace peptídico	23
1.1.1.2. Estructura Secundaria.....	24
1.1.1.3. Péptidos: estructura y función.	25
1.1.2 Carbohidratos	27
1.1.2.1. Conformación en carbohidratos	27
1.2. Interacciones proteína-carbohidrato	30
1.2.1. Interacciones de apilamiento (<i>stacking</i>) CH/ π	30
1.2.2. Enlaces de Hidrógeno.....	31
1.2.3. Interacciones electrostáticas	32
1.2.4. Implicación de las moléculas de agua e iones.....	33
1.3. Glicosidasas	34
1.3.1. Mecanismos de reacción	34
1.3.2. Inhibidores de glicosidasas.....	36
1.4. Resonancia Magnética Nuclear (RMN)	37
1.4.1. Aspectos generales de RMN	38
1.4.2. Relajación dipolo-dipolo y efecto Overhauser nuclear.....	40
1.4.3. Asignación de péptidos por RMN	43
1.5. RMN y reconocimiento molecular	45

1.5.1. Métodos basados en cambios en los movimientos rotacionales	47
1.5.1.1. Espectroscopia de diferencia de saturación (STD)	47
1.5.1.2. Secuencias WaterLOGSY	48
1.5.1.3. NOE-transferido (TR-NOE).....	48
1.5.1.4. Secuencias Carr-Purcell-Meiboom-Gill (CPMG).....	50
1.5.1.5. Experimentos de Relajación T1ρ	51
1.5.2. Métodos basados en cambios en la difusión translacional	52
Objetivos.....	53
2. Estudio estructural de análogos del péptido opiorfina.....	55
2.1. Barrido de alanina, <i>Alanine scanning</i> , en la opiorfina	57
2.2. Modulación de la actividad por sustitución del residuo de fenilalanina.....	57
2.3. Análisis conformacional por RMN.....	58
2.4. Conclusiones: relación estructura-actividad	67
2.5. Materiales y métodos experimentales.....	67
2.5.1. Síntesis de péptidos.....	67
1.5.2. Cálculo de actividades frente a neprilisina (NEP) y a aminopeptidasa-N (AP-N)	67
2.5.3. Resonancia Magnética Nuclear	68
2.5.4. Estructura tridimensional	68
3. Estudio estructural de análogos glicosilados de nociceptina.....	69
3.1. Actividad de los péptidos frente a los receptores de membrana NOP.....	70
3.2. Dicroísmo circular de la nociceptina y sus análogos glicosilados.....	71
3.3. Resonancia Magnética Nuclear de la nociceptina y sus análogos glicosilados.....	73
3.4. Conclusiones: Relación estructura-actividad.....	85
3.5. Materiales y métodos experimentales.....	85
3.5.1. Síntesis de péptidos.....	85
2.5.2. Cálculo de afinidades mediante ensayo competitivo frente a un radioligando.....	85

3.5.3. Helicidad mediante dicroísmo circular	86
3.5.4. Resonancia Magnética Nuclear	86
3.5.5. Estructura tridimensional	89
4. El uso de etiquetas difluoroacetamida como sonda para detectar interacciones carbohidrato-proteína.....	91
4.1. Análogos fluorados de quitobiosa y quitotriosa en estado libre	93
4.2. Análogos fluorados de quitobiosa y quitotriosa en estado asociado	100
4.3. Implicación del motivo difluoroacetamida en el proceso de reconocimiento	105
4.4. Reconocimiento molecular de los análogos fluorados de quitobiosa y quitotriosa.....	107
4.5. Conclusiones: Posibilidades de uso de esta sonda difluoroacetamida para detectar interacciones	110
4.6. Materiales y métodos experimentales.....	110
4.6.1. Resonancia Magnética Nuclear	110
4.6.2. Cálculos ab initio mediante el programa Gaussian 09.....	111
4.6.3. Dinámica molecular	111
5. Inhibidores de β -galactosidasa a través de modificaciones en anillo aglicónico	113
5.1. Conformación de los tiogalactósidos 1 y 2 en estado libre	116
5.2. Inhibidores de β -galactosidasa en estado asociado	123
5.3. <i>Docking</i> molecular de los inhibidores 1 y 2	127
5.4. Dinámica molecular de los complejos proteína-carbohidrato	128
5.5. Conclusiones: Desarrollo de inhibidores de glicosidasas mediante modificaciones en el anillo aglicónico.....	134
5.6. Materiales y métodos.....	135
5.6.1. Resonancia Magnética Nuclear	135
5.6.2. Cálculos de Mecánica Molecular	135
5.6.3. <i>Docking</i> molecular.....	136
5.6.4. Dinámica Molecular	136

Discusión	139
Conclusiones.....	143
Summary.....	147
Introduction	147
Chapter 1 – Structural analysis of opiorphin analogs by NMR.....	147
Chapter 2 – Structural analysis of glycopeptides of nociceptin by NMR	149
Chapter 3 – Difluoroacetamide moiety as probe to detect glycan-protein interactions by NMR	151
Chapter 4 – <i>E. coli</i> β -galactosidase inhibitors through modifications at the aglyconic moiety	153
Bibliografía.....	159

Abreviaturas

AMBER: Assisted Model Building with Energy Refinement

APN: AminoPeptidase N

BSA: Bovine Serum Albumin

CARA: Computer Aided Resonance Assignment

COSY: COReLation SpectroscopY

CSI: Chemical Shift Index

CYANA: Combined assignment and dYnamics Algorithm for NMR Applications

DC: Dichroism Circular

ΔG : variación de energía libre de Gibbs

DOSY: Diffusion-Ordered SpectroscopY

drNOP: *Danio rerio* NOP

E. coli: *Escherichia coli*

FRET: Fluorescence Resonance Energy Transfer

FSTD: ^{19}F -observed STD

Gal: Galactosa

GalNAc: N-Acetyl Galactosamine

GlcNAc: N-Acetyl Glucosamine

HEK293: Human Embryonic Kidney 293 cells

HMBC: Heteronuclear Multiple Bond Correlation

HPLC: High Performance Liquid Chromatography

HSQC: Heteronuclear Single-Quantum Coherence

IC50: Inhibitory Concentration 50

IPTG: IsoPropyl 1-ThioGalactopyranoside

J: scalar coupling constant

K_i : constante de inhibición

LigPrep: Ligand Preparation

MD: Molecular Dynamics

MM: Molecular Mechanics

NEP: Neutral EndoPeptidase

NeuNAc: N-Acetyl Neuraminic acid

NOE: Nuclear Overhauser Effect

NOESY: NOE SpectroscopY

NOP: NOcicePtin receptor

PDB: Protein Data Bank

RMN: Resonancia Magnética Nuclear

ROE: Rotating-frame Overhauser Effect

ROESY: ROE SpectroscopY

SDS: Sodium Dodecyl Sulfate

STD: Saturation Transfer Difference

STDreF: STD relay to ^{19}F

TFE: 2,2,2-TriFuoroEtanol

TLC: Thin Layer Chromatography

TOCSY: TOtal Correlation SpectroscopY

TR-NOESY: TRansferred NOE SpectroscopY

TR-ROESY: TRansferred ROE SpectroscopY

TSP: 2,2,3,3-tetradeuterio-3-(TrimethylSilyl)Propionic acid

WaterLOGSY: Water-Ligand Observed via Gradient SpectroscopY

WGA: Wheat Germ Agglutinin

τ_c : correlation time

Resumen

El reconocimiento molecular es el punto de partida de numerosos procesos biológicos. La comprensión del reconocimiento entre los ligandos y sus dianas es clave en la optimización y diseño de complejos de mayor afinidad. La combinación de datos de RMN con cálculos teóricos facilita desarrollar una aproximación a la comprensión de las bases moleculares de estos procesos de reconocimiento ligando-receptor. En esta tesis hemos estudiado cuatro sistemas ligando-receptor de interés biológico empleando esta estrategia combinando de RMN y modelización molecular (MM, MD y *ab initio*).

- Estudio estructural de análogos del péptido opiorfina.

Oporfina (QRFSR) es un péptido que impide la degradación de las encefalinas mediante la inhibición de las proteasas NEP y APN. Esta acción sobre el catabolismo de las encefalinas determina un efecto analgésico, demostrado en modelos animales. En este capítulo se han estudiado distintos análogos de la opiorfina cuya actividad ha sido correlacionada con los cambios estructurales observados en RMN.

El barrido de alanina en la secuencia de la opiorfina reveló que la sustitución del residuo Phe³ produce una pérdida drástica de la actividad frente a NEP y APN. Seguidamente, distintas modificaciones en esta posición mostraron la posibilidad de afectar la selectividad de la capacidad inhibitoria de este péptido. La sustitución de L-Phe por D-Phe mejoró la potencia de la opiorfina frente a APN, perdiendo su actividad frente a NEP.

Esta modulación de su actividad se correlaciona con cambios en los espectros de RMN de los péptidos que contienen el enantiómero D-Phe. La diferencia experimental más notable es el apantallamiento de los protones γ de la Arg² en los péptidos que contienen D-Phe. Los cálculos de MM mostraron la proximidad espacial entre el anillo aromático y la Arg², confirmándose el apantallamiento en los espectros RMN. Este estudio muestra una posible interacción catión/ π entre la D-Phe³ y la Arg² que favorece una restricción conformacional que determina una interacción selectiva con APN, mientras que la flexibilidad del péptido natural permite la inhibición dual, pero de menor afinidad frente a APN.

- Estudio estructural de análogos glicosilados de nociceptina

Nociceptina (FGGFTGARKSARKLANQ) es un neuropéptido que se relaciona con los opioides, pero no interacciona con los clásicos receptores opioides; su diana es el receptor NOP. Como su propio nombre indica, esta molécula está implicada en la modulación de la nocicepción. Este estudio demuestra que la glicosilación de este

péptido afecta a su actividad biológica, pudiéndose relacionar la modulación de la actividad con cambios en su estructura.

La actividad biológica de la opiorfina y tres análogos glicosilados fue ensayada mediante un ensayo por desplazamiento de radioligando. El análogo más potente fue la [Ser¹⁰-O- α -D-GalNAc]-nociceptina, de mayor afinidad que la nociceptina.

El dicroísmo circular no mostró diferencias entre las distintas moléculas. En cambio, RMN en presencia de micelas de SDS mostró diferencias en sus estructuras. Los cálculos mediante dinámica molecular restringida con datos experimentales de RMN (CYANA) mostraron que estos péptidos tienen tendencia α -hélice en presencia de SDS.

El péptido menos potente, cuya glicosilación es en Thr⁵, adopta una estructura plegada, aunque mantiene cierta tendencia de estructura helicoidal en el extremo C-terminal. En cambio, el glicopéptido más potente no interrumpe su estructura helicoidal, y muestra mayor tendencia de α -hélice que el péptido natural. Los datos estructurales de RMN indican que una mayor tendencia helicoidal de este péptido mejora su actividad frente a su diana.

- El uso de etiquetas difluoroacetamida como sonda para detectar interacciones carbohidrato-proteína

Las interacciones carbohidrato-proteína están presentes en gran número de procesos biológicos, pero la comprensión de estos procesos a escala molecular está limitada por la dificultad de obtener información a escala atómica de las moléculas de naturaleza glicídica. Así que es necesario el desarrollo de nuevas metodologías y estrategias que ayuden a la comprensión de estos procesos de reconocimiento molecular. La utilización de isótopos distintos al protón tiene beneficios técnicos para RMN, por lo tanto las sondas fluoradas son una herramienta útil en esta técnica. Este estudio propone la aplicación del grupo difluoroacetamido en el seguimiento de la interacción de dos moléculas que contienen residuos GlcNAc, *N,N'*-diacetil quitobiosa and *N,N',N''*-triacetil quitotriosa.

Esta nueva sonda de RMN fue evaluada frente a WGA mediante experimentos de RMN de ¹H y ¹⁹F. El proceso de reconocimiento de los ligandos analizados es semejante al observado en los análogos naturales. De hecho, el análogo fluorado de la quitobiosa interacciona por su extremo no-reductor y el análogo de la quitotriosa tiene dos modos de unión, WGA reconoce este análogo principalmente a través de su residuo central y su extremo no-reductor. Además, se evaluó mediante cálculos teóricos cómo afecta el grado de fluoración del grupo acetamido a la energía de interacción. Los grupos que contienen CF₂H o CFH₂ muestran una energía de unión más favorable cuando el enlace CH está orientada hacia el anillo aromático.

El grupo difluoracetamido ha mostrado ser una sonda de RMN útil en el análisis de la interacción de glicanos con el grupo NAc. Por lo tanto, esta sonda podrá ser utilizada en sistemas donde estén implicados residuos como GlcNAc, GalNAc o NeuNAc.

- Inhibidores de β -galactosidasa a través de modificaciones en anillo aglicónico

La ruptura catalítica del enlace glicosídico va precedida por la distorsión del sustrato tras su unión a la glicosidasa, estas enzimas reconocen los sustratos en conformaciones de alta energía en la trayectoria hacia el estado de transición. Este estudio investiga la influencia del motivo aglicónico en la afinidad de inhibidores de la β -galactosidasa de *E. coli*. Dos tiodisacáridos, *benzyl 3-deoxy-4-S-(β -D-galactopyranosyl)-4-thio- β -D-erythro-pentopyranoside* (**1**) and *benzyl 3-deoxy-4-S-(β -D-galactopyranosyl)-4-thio- β -D-threo-pentopyranoside* (**2**), que únicamente se diferencian en la estereoquímica en su posición C2 de la pentopiranos, han mostrado una diferencia significativa en su afinidad (Inhibidor **1** $K_i=800\mu\text{M}$, e inhibidor **2** $K_i=32\mu\text{M}$).

Estos dos inhibidores tienen un comportamiento conformacional distinto en su estado libre en disolución. **1** únicamente adopta confórmeros cuyo residuo aglicónico tiene una conformación silla 4C_1 , en cambio **2** presenta también otras geometrías con una silla 1C_4 .

La interacción de estos dos inhibidores a la β -galactosidasa de *E. coli* fue confirmada mediante experimentos STD y TR-NOESY. El análisis de los espectros TR-NOESY identificó los mismos confórmeros en estado unido para ambos inhibidores, $\text{syn}\phi_{1C4}/\text{syn}\psi_{1C4}$ y $\text{anti}\phi_{4C1}/\text{syn}\psi_{4C1}$. La β -galactosidasa selecciona o distorsiona estos inhibidores tanto modificando su torsión glicosídica ϕ como alterando la conformación silla de su motivo aglicónico. Además, las simulaciones de *docking* molecular y DM han mostrado estructuras de los complejos de esta proteína con sus inhibidores que explican los datos experimentales de RMN. Estos complejos fueron estables y los ligandos no difundieron al medio acuoso explícito manteniendo su conformación constante.

La conformación en silla 1C_4 del motivo aglicónico en estado libre establece una mejora en afinidad. La diferencia en la constante de inhibición en 25 veces demuestra que se puede modular la actividad de análogos de sustrato mediante las variaciones químicas y conformacionales en el anillo aglicónico.

Abstract

Molecular recognition is involved in numerous life processes. Understanding the variables which are involved in the interactions between biological targets and their ligands in these life processes has the potential to rationalize binding and selectivity features. A combined approach of NMR and theoretical MM and MD could clarify the system dynamics and provide deeper comprehension of the molecular basis of ligand-receptor recognition. In this thesis the analysis of four different ligand-receptor systems of biological relevance has been accomplished by this combined approach.

- Structural analysis of opiorphin analogs by NMR

Opiorphin (QRFSR) slows down enkephalin degradation by inhibiting both NEP and APN. This peptide produces antinociceptive effects as it was demonstrated in rodent models. In this thesis different opiorphin analogs were characterized by their activity against NEP and APN, and the most remarkable activity findings were analyzed in structural context by a combined approach NMR and MM.

The activity of these analogs was measured versus NEP and APN by fluorimetric assays. Alanine scanning showed Phe³ should play a relevant role in the recognition. Subsequently, distinct modifications at Phe³ modulate its dual inhibition. D-Phe³-containing peptides were one order of magnitude more potent than the parent peptide in APN assays, but inactive in NEP assays.

NMR corroborates the different tendencies between L-Phe³-containing and D-Phe³-containing peptides. The most remarkable experimental difference is severe shielding of Arg² H γ 2 and H γ 3 protons to higher field in D-Phe³-containing peptides. MM calculations confirmed the spatial proximity between the aromatic ring and the H γ protons of Arg², explaining the shielding at higher field observed by NMR. In contrast, flexible native opiorphin cannot be explained with a single conformation. In this study a higher-affinity selective APN inhibitor has been identified, whose potency enhancement depends on its conformational rearrangement in solution.

- Structural analysis of glycopeptides of nociceptin by NMR

Nociceptin (FGGFTGARKSARKLANQ) is a neuropeptide which is related to opioids. However, nociceptin doesn't interact with the classical opioid receptors; it is recognized by the N/OFQ receptor (NOP). As its name indicates, early studies showed its involvement in the pain modulation. In this chapter distinct glycosylations at different parts of nociceptin were studied and its biological activity was correlated with structural changes.

The biological activity by a competition radioligand binding assay showed that the most potent peptide is [Ser¹⁰-O- α -D-GalNAc]-nociceptin, even slightly higher affinity than

parent peptide. On the contrary, affinity of [Thr⁵-O- α -D-GalNAc]-nociceptin is one order of magnitude less active than nociceptin.

Far-UV CD didn't find differences which can be correlated with changes in their inhibition constants. However, NMR data in the presence of SDS micelles showed significant differences. NMR data show that these peptides have the tendency to form α -helix structures in SDS micelles, and restricted molecular dynamics calculation including information from NOESY spectral data (CYANA) provided a more detailed structural interpretation.

3D structures of the four peptides were obtained and compared with their inhibition constants. The least potent peptide, whose glycosylation is on Thr⁵, showed a bend conformation. Nevertheless, the most potent glycopeptide does not disrupt the helical structure of nociceptin, and shows even higher α -helix tendency than the parent peptide. Thus, glycosylation of nociceptin modulates the binding affinity to its receptor. NMR results conclude that a higher α -helix tendency correlate with an enhancement binding affinity.

- Difluoroacetamide moiety as probe to detect glycan-protein interactions by NMR

The development of new protocols and technologies is important in order to gain insights into those molecular recognition processes where carbohydrate structures are involved. In NMR, the use of isotopes other than ¹H shows several benefits from a technical point of view. This study proposes the application of a difluoroacetamide moiety, that allows to get NMR information from both isotopes ¹⁹F and ¹H, as a sensor to monitor the binding of two GlcNAc-containing molecules, *N,N'*-diacetyl chitobiose and *N,N',N''*-triacetyl chitotriose.

This new NMR probe was evaluated with two analogues against WGA, and their binding was monitored by distinct ¹H-NMR and ¹⁹F-NMR experiments. Besides, it was clearly demonstrated the existence of a molecular recognition process in which there is a preferred interaction with the residue at the non-reducing end for chitobiose analogue, and with the central and non-reducing end residues for chitotriose analogue. A molecular modeling protocol confirms that the molecular recognition features of both complexes are the same as for the parent compounds. Besides, theoretical calculations on simplified models confirm that the interaction energy depends on the degree of fluorination of the acetamide group. The ones containing CF₂H and CFH₂ groups showed the lowest interaction energy when it is the CH bond the one oriented towards the aromatic ring.

Thus, difluoroacetamide group is a practical tagging method for analyzing the interaction of NAc-containing glycans against receptors by NMR. This probe can be utilized in interacting systems where residues, such as GlcNAc, GalNAc or NeuNAc, are involved.

- *E. coli* β -galactosidase inhibitors through modifications at the aglyconic moiety

The catalytic cleavage of a glycosidic linkage is preceded by a molecular distortion of the substrate upon binding to glycosidases. In fact, these enzymes recognize their substrate in high-energy conformations in the trajectory towards the transition state. *E. coli* β -galactosidase is considered as a model protein of glycosidases, whose mechanism has been deeply investigated. Herein, it was investigated how affinity of two β -galactosidase inhibitors can be modulated through conformational preferences at their aglyconic moiety. These thiodisaccharides, *benzyl 3-deoxy-4-S-(β -D-galactopyranosyl)-4-thio- β -D-erythro-pentopyranoside* (**1**) and *benzyl 3-deoxy-4-S-(β -D-galactopyranosyl)-4-thio- β -D-threo-pentopyranoside* (**2**), differ only in the stereochemistry at position C2 of the pentopyranoside at the reducing-end, but the difference in their inhibition constants was reported to be 25-fold (Inhibitor 1 $K_I=800\mu\text{M}$, and Inhibitor 2 $K_I=32\mu\text{M}$).

Conformational studies in free state concluded that **1** adopts only conformers with the aglycone in the 4C_1 conformation; and the conformational ensemble of **2** shows conformers where the aglycone adopts both chair geometries, 4C_1 and 1C_4 .

The molecular recognition of both inhibitors by *E. coli* β -galactosidase was confirmed by STD and TR-NOESY experiments. TR-NOESY experiments determine that bound conformers of both inhibitors are $\text{syn}\phi_{1C4}/\text{syn}\psi_{1C4}$ and $\text{anti}\phi_{4C1}/\text{syn}\psi_{4C1}$. Thus, β -galactosidase is selecting a distorted conformation different from the mayor one in the free state, either by turning from $\text{syn}\phi$ to $\text{anti}\phi$ at the glycosidic bond, or by changing from a 4C_1 to a 1C_4 chair at the aglycone moiety. Afterward, molecular docking and MD simulations were performed in order to provide plausible three-dimensional structures for the complexes which could explain the NMR experimental data.

The conformational analysis of these two inhibitors explains the difference between their inhibitory activities. A distorted geometry at aglyconic moiety in free state determines an enhancement of inhibition activity. The differences in the inhibition of these molecules show the interest of a new generation of molecules which can modulate their activity by modifying their aglyconic moiety at selected positions.

1. RMN y Reconocimiento Molecular

El reconocimiento molecular es clave en numerosos procesos biológicos. La comprensión de los procesos de reconocimiento entre los ligandos y sus dianas es clave en la optimización y diseño de nuevos ligandos con afinidades mejoradas. La Resonancia Magnética Nuclear, RMN, permite identificar qué compuestos interaccionan con una macromolécula dada, qué conformación adoptan las moléculas en el centro de reconocimiento y qué regiones de los compuestos están implicadas en su unión. Los métodos de RMN se dividen en dos categorías, experimentos desde el punto de vista del receptor o desde la perspectiva del ligando.

En esta tesis hemos estudiado distintos sistemas de interés biológico, en los que están implicados bien péptidos o carbohidratos, empleando métodos de RMN desde la perspectiva del ligando. No obstante, antes de describir los distintos aspectos metodológicos, vamos a describir brevemente algunas características básicas estructurales y conformacionales de los dos tipos de moléculas que hemos empleado en esta tesis, las características globales de sus interacciones, así como una breve descripción de los métodos generales de RMN. Hay que mencionar en cualquier caso que la aplicación de la RMN en el análisis estructural y conformacional de biomoléculas, especialmente péptidos, está ampliamente extendida, y hoy en día se aplica de manera habitual en muchos laboratorios de España y del mundo¹⁻⁸.

1.1. Péptidos y Carbohidratos

1.1.1. Péptidos

Se denomina péptido a la cadena lineal de aminoácidos de una determinada longitud unidos mediante enlaces peptídicos. El aminoácido es la unidad estructural de proteínas y de péptidos, y está caracterizado por un grupo amino y un grupo carboxílico anclados a su carbono alfa, $C\alpha$. Los otros dos sustituyentes de este $C\alpha$ son un hidrógeno, *hidrogeno alfa*, y otro sustituyente que determinara la naturaleza de dicho aminoácido, *cadena lateral*. Al ser los cuatro sustituyentes distintos en el $C\alpha$, estos compuestos presentan isomería óptica. De hecho, exceptuando la glicina, el resto de los 20 residuos naturales provenientes de la expresión ribosomal son L-aminoácidos.

Los 20 aminoácidos naturales se clasifican según la naturaleza y propiedades fisicoquímicas de la cadena lateral, y se pueden nombrar por dos sistemas de abreviaturas, con una única letra o con códigos de tres letras. Según esta clasificación tendremos cuatro principales tipos de aminoácidos: aminoácidos apolares, aminoácidos polares neutros, aminoácidos cargados positivamente y los aminoácidos cargados negativamente (Figura 1).

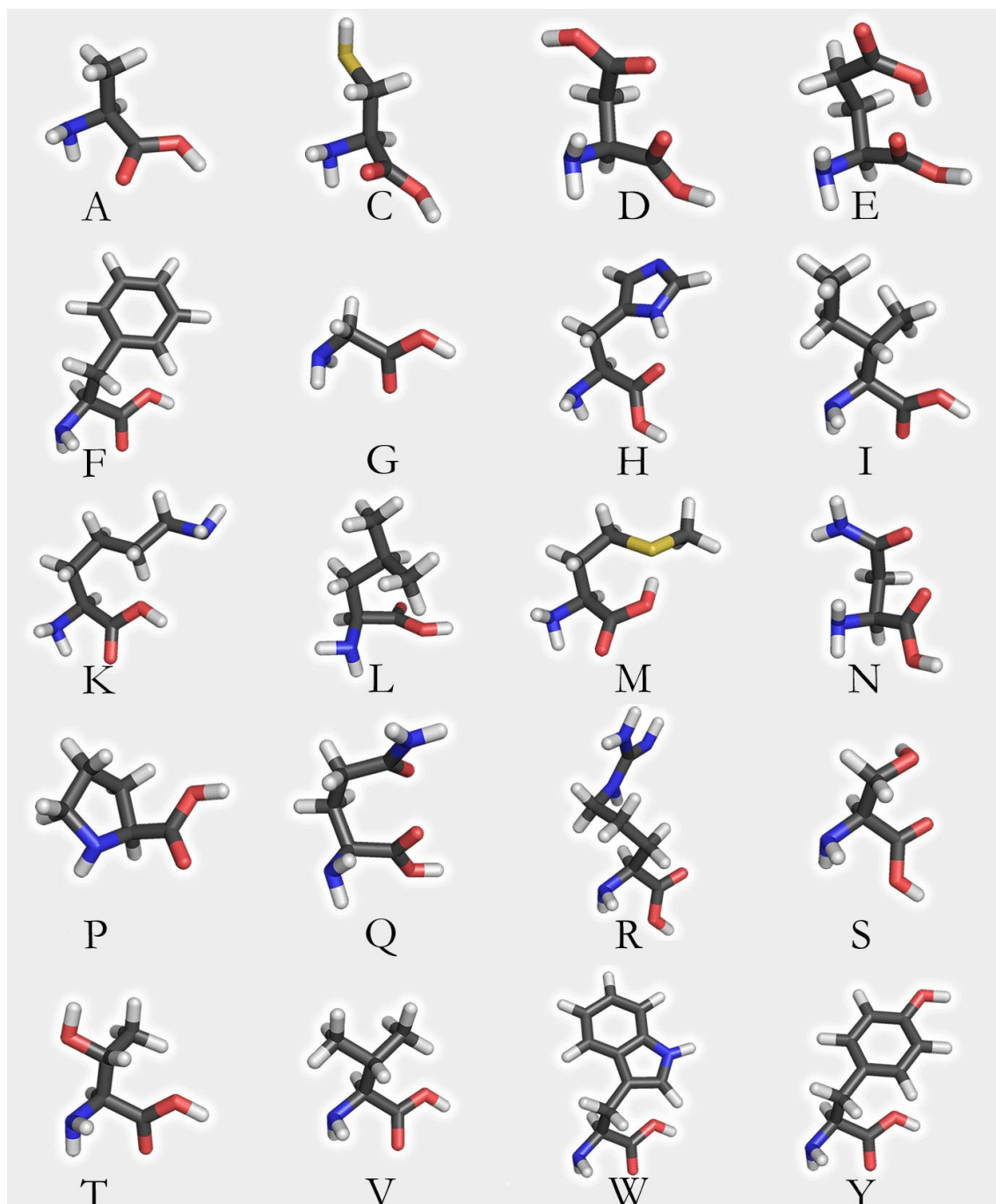


Figura 1: Representación tridimensional de los 20 aminoácidos naturales. Los aminoácidos han sido identificados con la abreviatura de una letra. Los aminoácidos se clasifican en cuatro principales tipos: aminoácidos apolares (Alanina[A;Ala], Fenilalanina[F;Phe], Glicina[G;Gly], Isoleucina[I;Ile], Leucina[L;Leu], Metionina[M;Met], Prolina[P;Pro], Valina[V;Val], Triptófano[W;Trp] y Tirosina[Y;Tyr]); aminoácidos polares neutros (Cisteína[C;Cys], Asparagina [N;Asn], Glutamina[Q;Gln], Serina [S;Ser] y Treonina[T;Thr]); aminoácidos cargados positivamente (Histidina[H;His], Lisina[K;Lys] y Arginina[R;Arg]); y los aminoácidos cargados negativamente (Ácido Aspárticos[D;Asp] y Ácido Glutámico[E;Glu]).

1.1.1.1. Enlace peptídico

El enlace peptídico es la condensación de un grupo amino de un aminoácido con un grupo carboxilo de otro, tras la pérdida de una molécula de agua, produciendo la formación de un enlace covalente tipo amida. Este enlace determina un serie de limitaciones conformacionales debido a la resonancia entre la forma neutra de enlace sencillo C-N y la forma de enlace doble C=N. Esto determina que el enlace peptídico tenga un carácter parcial de doble enlace, con una distancia de 1,32 Å entre el carbono carbonílico de un aminoácido y el nitrógeno amínico del siguiente aminoácido. De hecho, se determina un plano peptídico con seis átomos coplanarios cuyos sustituyentes tienen una disposición *trans* o *cis* entre los C α de ambos aminoácidos unidos, según el ángulo diedro ω , sobre el enlace C-N, sea de 180° o 0° respectivamente. En general, la disposición *cis* está desfavorecida energéticamente, por lo tanto la gran mayoría de los enlaces peptídicos presentan una disposición *trans* (Figura 2). Una excepción es la prolina que, debido a un mayor impedimento estérico de la disposición *trans*, adquiere más fácilmente una disposición *cis*.

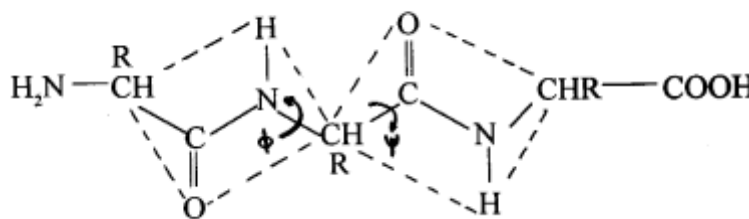


Figura 2: Representación bidimensional del enlace peptídico. Se ha representado el carácter parcial de doble enlace mediante el plano peptídico (línea discontinua) y los dos torsionales con libertad de giro, ϕ y ψ .

La flexibilidad de las cadenas peptídicas es dependiente del esqueleto de la cadena y de los distintos enlaces de sus cadenas laterales. En el esqueleto peptídico hay tres enlaces, y consecuentemente tres torsiones, a tener en cuenta. La torsión del plano peptídico ω , respecto al enlace CO-NH, adopta únicamente dos valores como hemos mencionado anteriormente. Las otras torsiones son los ángulos denominados ϕ , enlace NH-C α , y ψ , enlace C α -CO. Estos dos torsionales son enlaces sencillos, así que no tienen restricciones como en el caso de ω . En cambio, todos los valores de ϕ y ψ no están permitidos debido a la posibilidad de choques estéricos para algunos valores. La representación bidimensional de las regiones toleradas dependiente de ambos ángulos ϕ y ψ es lo que se denomina *mapa de Ramachandran* (Figura 3). Este mapa es similar para todos los aminoácidos, excepto para el aminoácido glicina. El valor de estos ángulos va a definir la estructura secundaria local de la cadena peptídica para ese aminoácido. En las cadenas laterales sus torsionales se denominan con la letra griega χ y con un subíndice numérico dependiendo de la distancia en enlaces al C α . Así que χ_1

será la torsión entre el C_α y el C_β . Sus sustituyentes tienen más tendencia a adoptar una disposición *alternada*, cuyo valor dependerá del tipo de aminoácido y del valor de los ángulos ϕ y ψ del esqueleto peptídico⁹⁻¹⁰.

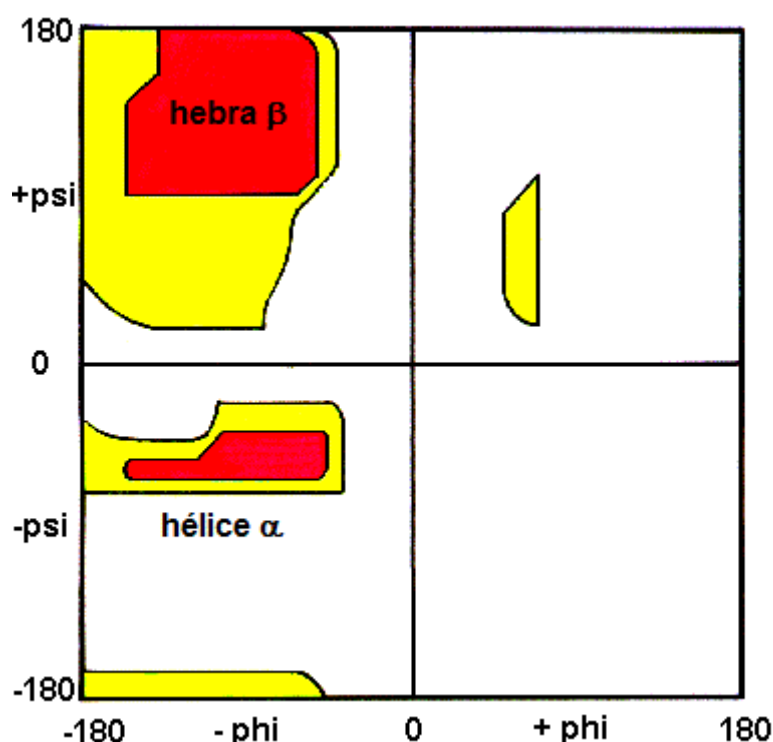


Figura 3: Mapa de *Ramachandran* en el que se representa el giro rotacional del torsional ϕ (phi) en el eje de abscisas y el del torsional ψ (psi) en el eje de ordenadas.

1.1.1.2. Estructura Secundaria

El diagrama de *Ramachandran* muestra principalmente dos zonas de ángulos ϕ y ψ permitidos. Estas zonas están asociadas con dos geometrías de estructura secundaria, la hélice α ($\psi \sim -70^\circ - 0^\circ$) y la hebra β ($\psi \sim +100^\circ - +180^\circ$). La sucesión de, al menos, cinco aminoácidos consecutivos cuyo ángulos ψ están en torno a los -60° determina una estructura de hélice α tipo *R* (dextrógira). Este tipo de hélice presenta enlaces de hidrogeno entre el grupo CO de un aminoácido (*i*) y el NH del aminoácido cuya posición es *i*+4, determinando que una vuelta de hélice α se dé cada 3,6 aminoácidos y avance 5,4 Å cada vuelta. Su longitud varía según si se encuentran en proteínas globulares, cuya longitud media es 12 aminoácidos, o en proteínas de membrana, cuya longitud es mayor para atravesar la membrana lipídica. Las cadenas laterales de los aminoácidos quedan expuestas de manera radial, ligeramente dirigidos al extremo amino terminal. Las características de las cadenas laterales de los aminoácidos que forman la hélice determinarán su estabilidad y su naturaleza (hélice anfipática, hélice apolar o hélice polar). Hay otras estructuras helicoidales pero tienen otras características y otro patrón de enlaces de hidrógeno¹¹.

La presencia de, al menos, dos aminoácidos consecutivos con valores ψ entre $+100^\circ$ y $+180^\circ$ se denomina hebra β . Su conformación extendida determina un avance por cada residuo $3,5\text{\AA}$, frente a los $1,5\text{\AA}$ de la hélice α . Este carácter extendido determina la disposición alternada de los grupos CO y NH de la cadena, lo que facilita la asociación de varias hebras generando una *lámina β* . Si las hebras de la lámina β poseen el mismo sentido bioquímico se denomina lámina paralela, si el sentido bioquímico es alternado se llama lámina antiparalela, y, por último, las lámina mixtas tienen hebras dispuestas paralelamente y antiparalelamente. Estas distintas estructuras de lámina β colocan las cadenas laterales en ambas caras del plano de manera alternada.

No todos los aminoácidos de una cadena peptídica participan en estos tipos de estructuras secundarias. En ciertas regiones, los residuos adoptan otras estructuras más o menos definidas que conectan las anteriores estructuras secundarias; por ejemplo, los bucles y los giros. En estas regiones de la cadena abundan los residuos de glicina, cuya libertad conformacional es mayor en torno a los ángulos ϕ y ψ .¹¹

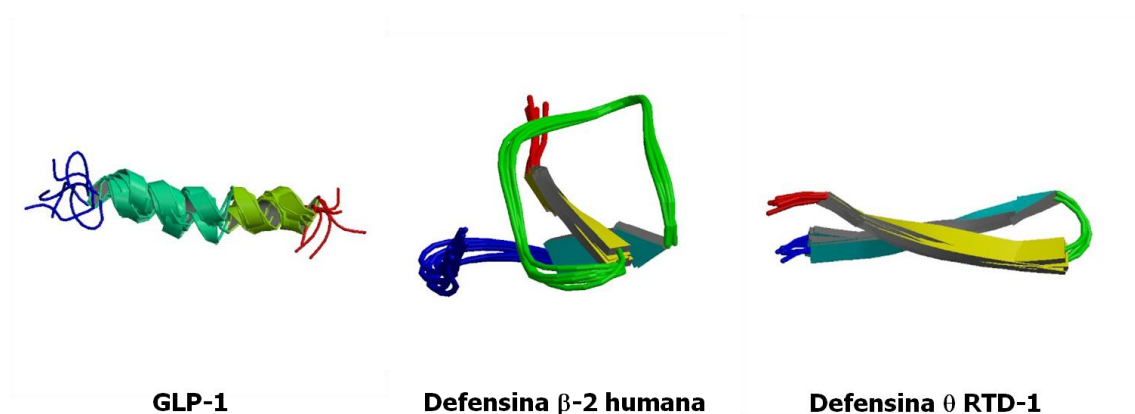


Figura 4: Estructuras tridimensional de péptidos obtenidos mediante técnicas de RMN. Como ejemplo de hélice α se muestra el péptido GLP-1 (*Glucagon-like Peptide-1-(7-36)-amide*)¹². Se muestran dos ejemplos de hebra β , la estructura de péptido defensina β -2 humana¹³ y la estructura de la defensina θ (RTD-1)¹⁴.

1.1.1.3. Péptidos: estructura y función.

Un péptido es una cadena corta de aminoácidos, y se diferencia únicamente de una proteína por la longitud de la cadena peptídica. Por lo tanto, un péptido es una cadena corta de dos o más aminoácidos, y una proteína es una cadena larga que está formada por una o más cadenas polipeptídicas. En la célula los péptidos están encargados de múltiples funciones biológicas, incluso, algunos péptidos actúan como hormonas siendo liberadas por una célula y actuando sobre otras en otra parte del cuerpo¹⁵.

La insulina fue la primera molécula de naturaleza proteica que fue secuenciada. *Frederick Sanger* determinó la secuencia de esta hormona y demostró que su estructura

primaria era lineal no ramificada¹⁶⁻²¹. La insulina es activa en su forma monomérica y está compuesta por dos cadenas; sus cadenas A y B tienen 21 y 30 aminoácidos respectivamente. Tiene tres puentes disulfuro cuyas posiciones están conservadas en la clase *Mammalia*; un puente disulfuro une dos cisteínas de la cadena A y los otros dos unen covalentemente la cadena A y B. Su estructura tridimensional consta de tres hélices, 2 hélices en la cadena A y otra en la B. La estabilidad que confieren las uniones entre sus cisteínas es complementada por un núcleo de residuos hidrofóbicos²². El análisis del complejo entre la insulina y su receptor demostró que la insulina interacciona a través de dos áreas con la estructura dimérica de su receptor²³⁻²⁴. Otro ejemplo de hélice α es el glucagón. Esta hormona peptídica, que tiene una acción fisiológica opuesta al efecto de la insulina, adopta una estructura secundaria de α hélice en gran parte de sus 29 aminoácidos²⁵⁻²⁷.

Como ya se ha mencionado, en la naturaleza las cadenas polipeptídicas pueden adoptar otras estructuras distintas de hélice α , por ejemplo péptidos como las defensinas presentan estructuras de hebra β (Figura 4). Estos péptidos son moléculas anfipáticas con capacidad antimicrobiana cuya rígida y compacta estructura secundaria, que está estabilizada por 3 o 4 puentes disulfuro, con gran resistencia a la degradación y digestión por peptidasas. Ambos tipos de estructuras secundarias no son excluyentes y coexisten en el mismo péptido, así por ejemplo en algunas defensinas de insectos²⁸ y de plantas²⁹ se dan motivos de hélice α .

Además de la estructura de hélice α y la de hebra β , los péptidos puedan adoptar conformaciones muy diferentes. Por ejemplo son capaces de adoptar uno de los motivos más comunes en una proteína, el giro β ³⁰. Este motivo está presente en muchas regiones de contacto proteína-proteína, por lo tanto péptidos que mimeticen estas estructuras serán una estrategia para el diseño de inhibidores de estas interacciones entre proteínas³¹⁻³². Otras estructuras de péptidos están determinadas por la posición conservada de residuos de cisteína en su secuencia. En el tracto intestinal se han identificado unos péptidos denominados *trébol* (*trefoil peptides*) por la estructura que adquiere su secuencia de 38-39 aminoácidos. La secuencia consenso (C-X₉₋₁₀-C-X₉-C-X₄-C-C-X₁₀-C) tiene tres puentes de disulfuro cuya configuración está conservada en estos péptidos, y genera una estructura con tres bucles (*loops*) que da nombre a la familia³³⁻³⁴. Los lantibióticos son otro ejemplo de estructuras peptídicas. En este caso, los péptidos incluyen aminoácidos no proteinogénicos o post-translacionalmente modificados con uniones covalentes entre las cadenas laterales de los aminoácidos generando estructuras cíclicas. Estas bacteriocinas poseen unas estructuras con múltiples ciclos que generan una estructura rígida con resistencia a la degradación proteolítica y a la inactivación térmica³⁵⁻³⁶. Todos estos ejemplos muestran la diversidad estructural de los péptidos y denotan la implicación de su estructura en la función que realizan. La estructura que adopta la cadena lineal de aminoácidos en el espacio se puede relacionar con su actividad hormonal, estructural o antimicrobiana.

1.1.2 Carbohidratos

Los carbohidratos (glicanos, azúcares, oligosacáridos) están implicados en todos los procesos biológicos que se llevan a cabo en una célula, son utilizadas como fuente de energía y realizan funciones estructurales y de reconocimiento. Su unidad básica son los monosacáridos. Los carbohidratos más complejos están formados por estos monómeros unidos mediante enlaces glicosídicos, pudiendo alcanzar cadenas de gran tamaño, lineales y ramificadas. El enlace O-glicosídico entre sucesivos residuos se produce por condensación y deshidratación del grupo hidroxilo hemiacetálico de un monosacárido con uno de los hidroxilos de otro monosacárido. Esta unión covalente entre dos monosacáridos genera dos ángulos torsionales denominados ϕ ($H_1^{(i)}-C_1^{(i)}-O_1^{(i)}-C_X^{(i-1)}$) y ψ ($C_1^{(i)}-O_1^{(i)}-C_X^{(i-1)}-H_X^{(i-1)}$). Si el enlace glicosídico implica a un carbono que no está localizado en el anillo, la posición relativa entre los dos residuos también estará definida por el torsional ω ($O_1^{(i)}-C_X^{(i-1)}-C_Y^{(i-1)}-H_Y^{(i-1)}$) (Figura 5).

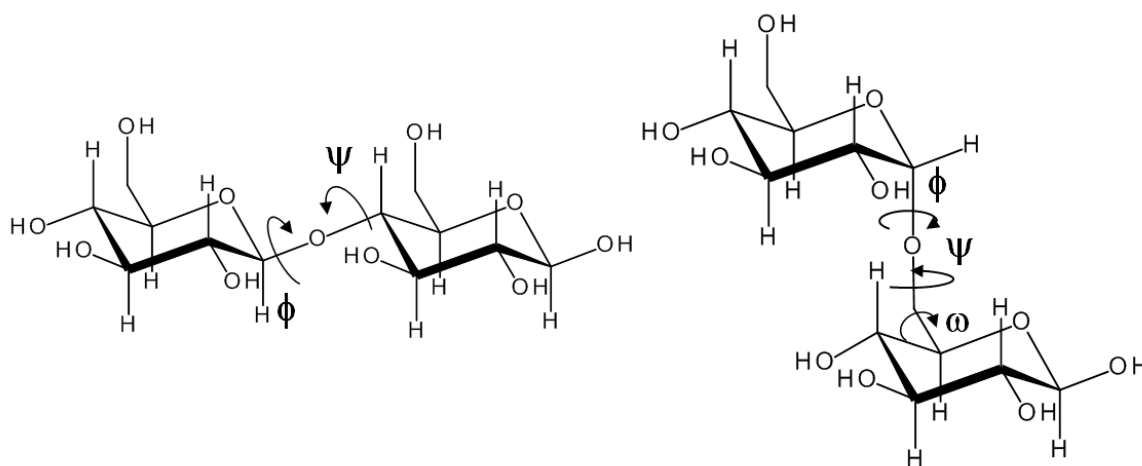


Figura 5: Enlace glicosídico y sus ángulos torsionales. *Izda.*, la celobiosa está formada por dos moléculas de glucosa unidas por una unión $\beta(1\rightarrow4)$; *dcha.*, la isomaltosa consiste en dos residuos de glucosa con una unión $\alpha(1\rightarrow6)$.

1.1.2.1. Conformación en carbohidratos

Un monosacárido puede representarse linealmente mediante una proyección de Fischer. Sin embargo, los monosacáridos se encuentran en la naturaleza formando hemiacetales cíclicos, denominándose *furanosa* si el anillo está formado por 5 átomos y *piranosa* si el anillo es de 6 átomos. Además, la ciclación de estas moléculas determina la aparición de un nuevo centro quiral, el *centro anomérico*. Si el extremo reductor no está bloqueado, se produce la interconversión del anómero α en su otro anómero, β , mediante un proceso que se denomina *mutarotación*. La disposición de los sustituyentes en este carbono está modulada por los efectos anomérico y exo-anomérico. El *efecto anomérico*³⁷⁻³⁸ determina que los sustituyentes electronegativos adopten prioritariamente una orientación axial frente a una ecuatorial, debido a la deslocalización hiperconjugativa de uno de los pares de electróns no enlazantes del

oxígeno endocíclico con el orbital σ^* del enlace entre el carbono anomérico y el oxígeno del grupo hidroxilo. Este efecto estabiliza la configuración α en la serie D de los monosacáridos, contraponiéndose a lo predecible según efectos estéricos. Por otro lado, el *efecto exo-anomérico*³⁹ favorece las conformaciones cuyo ángulo ϕ permita la interacción entre un par de electrones no compartidos del oxígeno exocíclico y el orbital σ^* del enlace endocíclico C₁-O₅. Esto favorece las conformaciones EXO y EXO-ANTI (Figura 6).

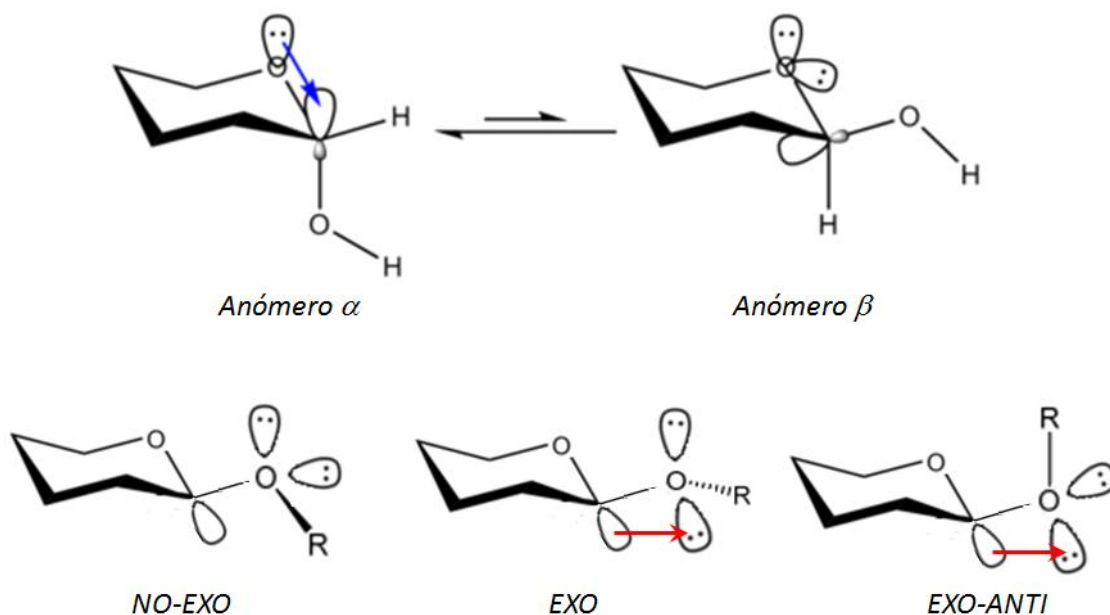
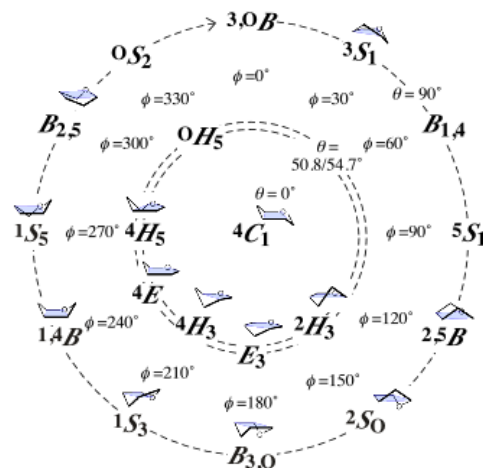



Figura 6: Representación esquemática de la conjugación de orbitales en el efecto anomérico y en el efecto exo-anomérico.⁴⁰

Además, el anillo adopta distintas conformaciones que se definen según el número de átomos que estén alineados en el plano del ciclo. En un anillo de piranosa, los 6 átomos del ciclo no se colocan en el mismo plano y disponen uno o dos de estos átomos arriba o abajo del plano generado por el resto de átomos. Según la disposición de sus átomos se diferencian 5 conformaciones: silla o *Chair* (C), bote o *Boat* (B), bote torcido o *Skew* (S), semi-silla o *Half-chair* (H) y sobre o *Envelope* (E) (Figura 7). La conformación o conformaciones que adopte una piranosa depende de sus sustituyentes. Normalmente todo sustituyente voluminoso se coloca en posición ecuatorial, evitando interacciones 1,3 diaxiales. De este modo, los anillos de piranosa de la D-glucosa, la D-galactosa o la D-manosa tienden a adoptar una conformación de silla ⁴C₁, colocando la mayoría de sus grupos hidroxilo ecuatorialmente. Asimismo, el grupo hidroximetilo en posición C₆ adopta tres posibles rotámeros en torno a su enlace C₅-C₆. La distribución poblacional depende de la configuración del hidroxilo en posición C₄. Si el hidroxilo en C₄ es ecuatorial (configuración *gluco*) el grupo hidroxilo de C₆ toma una orientación *gauche-gauche* (GG) y *gauche-trans* (GT), mientras que la disposición axial en C₄





α -D-glucosa β -D-galactosa L-fucosa ácido α -L-idurônico

TG *GG* *GT*

- 29 -

1.2. Interacciones proteína-carbohidrato

El estudio de estas interacciones a nivel atómico es crucial tanto para la comprensión de su papel biológico, así como para el diseño de moduladores de esta interacción⁴⁴⁻⁴⁷. A nivel intracelular, estas interacciones tienen un papel fundamental en el tráfico vesicular⁴⁸. Por otra parte, las membranas celulares presentan al exterior diversas proteínas receptoras de carbohidratos expuestos por otras células o presentes en el suero. Así, por ejemplo, las selectinas presentes en células del epitelio vascular actúan en la movilización de neutrófilos y monocitos a la zona dañada⁴⁹ reconociendo carbohidratos en la superficie de estas células. Otro caso ampliamente estudiado es la heparina, este polisacárido está implicado en el proceso de coagulación sanguínea al ser reconocido por diversos factores de coagulación⁵⁰.

Todos estos fenómenos están mediados por los múltiples grupos funcionales de los carbohidratos y su presentación tridimensional a sus receptores dando lugar a procesos de reconocimiento molecular exquisitamente específicos. Los carbohidratos son moléculas anfifílicas. Por tanto, distintos tipos de interacciones regulan el proceso de reconocimiento. Sus grupos polares están involucrados como aceptores y donadores de enlaces de hidrógeno, tanto con el esqueleto peptídico de la proteína como con las cadenas laterales polares. Las regiones apolares de los carbohidratos tienden a aproximarse a los residuos aromáticos del sitio de unión⁵¹⁻⁵². De hecho, la interacción entre los hidrógenos de la *cara* apolar y los residuos aromáticos del receptor es dependiente del tamaño y de la riqueza en electrones del anillo aromático⁵³. Además, se ha propuesto que la desolvatación de las zonas apolares de estos ligandos promueve la interacción entre las moléculas de estos complejos⁵⁴. Estas interacciones pueden ser claramente observadas en un vasto número de estructuras de estos complejos (base de datos Glyco3d, <http://glyco3d.cermav.cnrs.fr>).

1.2.1. Interacciones de apilamiento (*stacking*) CH/ π

Estas interacciones tienen lugar entre los residuos aromáticos del receptor y las caras apolares de los azúcares. En estos contactos están implicados normalmente tres grupos CH de un anillo de piranosa y la nube electrónica π de un residuo aromático, denominándose interacciones *stacking* CH/ π ⁵⁵. Los cálculos de mecánica cuántica en fase gaseosa establecen que la energía de interacción entre un residuo aromático y un monosacárido está en el rango de 3-6 kcal/mol. Sin embargo, medidas experimentales en agua muestran una menor contribución energética entre un monosacárido y un anillo aromático, de aproximadamente 1.5 kcal/mol. Estas discrepancias entre los valores teóricos y los experimentales remarcan que estas interacciones están moduladas por aspectos entrópicos y efectos del disolvente⁵⁵. El residuo aromático puede interaccionar a través de ambos lados del plano. Sin embargo, la orientación de un monosacárido frente al anillo aromático dependerá de las uniones CH que estén interaccionando con la nube electrónica π . Por ejemplo, la α -D-Galactosa interacciona preferentemente a través

de su cara *beta*. Otros monosacáridos, como la β -D-Glucosa, β -D-xilosa, β -D-NAcGlucosamina, pueden interaccionar a través de ambas.

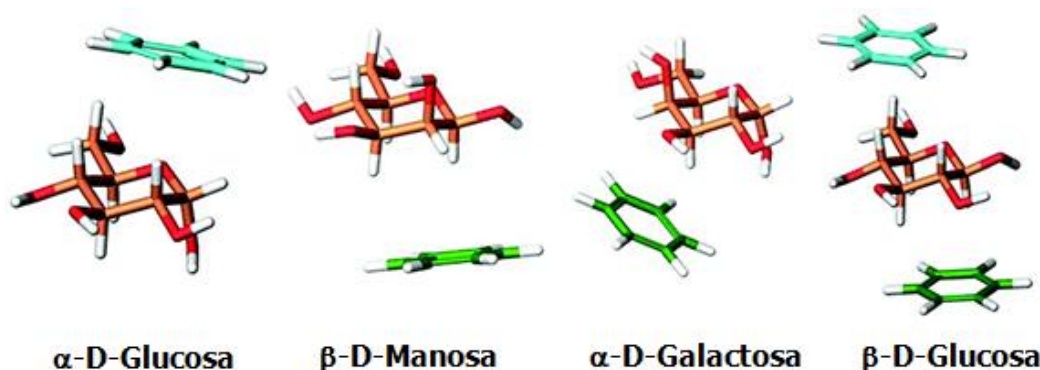


Figura 10: Representación esquemática de las interacciones *stacking* CH- π de α -D-Glucosa, β -D-Manosa, α -D-Galactosa, β -D-Glucosa.⁵⁵

Además de las uniones CH, el grupo metilo de grupos acetamido de los motivos GlcNAc, GalNAc y NeuNAc tienen contactos CH/ π con residuos aromáticos ayudando al reconocimiento de los ligandos. Estas interacciones se han observado en el reconocimiento de la aglutinina del germen de trigo⁵⁶⁻⁵⁸ o de la hemaglutinina del virus de la gripe⁵⁹.

1.2.2. Enlaces de Hidrógeno

Los enlaces de hidrógeno son un componente principal del proceso de reconocimiento molecular de carbohidratos. Los grupos hidroxilo tienen un papel doble en este tipo de interacciones, actuando como aceptores y donadores. Cada hidroxilo puede actuar como aceptor de dos uniones y como donador en una interacción, ya que los dos pares de electrones no compartidos aceptan su unión con dos hidrógenos y el protón de este hidroxilo interaccionará con un átomo electronegativo⁶⁰⁻⁶¹.

En el centro de reconocimiento de las proteínas, distintos aminoácidos forman enlaces de hidrógeno con estos hidroxilos. Por ejemplo, un residuo ácido puede actuar como aceptor de uno o dos hidroxilos. En cambio, otros grupos como los protones amida de la cadena principal o las cadenas laterales de Asn o Gln actúan como donadores. Otros aminoácidos de la proteína (Ser, Thr y Tyr) podrán tener un carácter ambivalente en el reconocimiento, como aceptores o donadores. La formación de estas interacciones con el receptor restringe la libertad torsional de estos grupos hidroxilo, determinando un coste entrópico en el procesos de unión, que se contrapone al componente entálpico de estas uniones intermoleculares. De hecho, esta penalización entrópica se refleja en la baja frecuencia con la que los grupos hidroxilos de proteína están implicados en puentes de hidrógeno⁶².

El oxígeno del anillo heterocíclico también puede actuar como aceptor a través de sus dos pares de electrones no compartidos. Sus interacciones no implican selectividad, dada su posición conservada en todos los anillos de piranosa. Otro grupo implicado en estas interacciones es el motivo acetamido de residuos como GlcNAc, GalNAc y NeuNAc. Este motivo, de geometría plana, es capaz de interaccionar como donador de un enlace de hidrógeno a través de su grupo amida y usar el oxígeno como aceptor⁶³⁻⁶⁴.

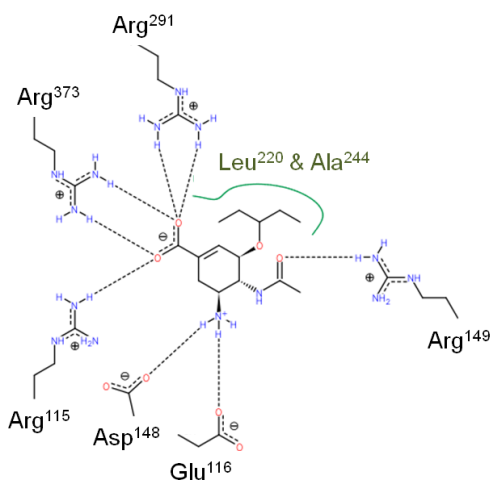
La red de enlaces de hidrógeno determina un reconocimiento estereoespecífico. La estereoespecificidad se consigue por la complementariedad entre los residuos del sitio de unión y los grupos donadores o aceptores del carbohidrato. En el sitio de unión, la mayoría de los residuos tienen una geometría plana, con varios grupos funcionales que permiten múltiples enlaces de hidrógeno, pudiendo incluso formar enlaces de hidrógeno bifurcados. La bifurcación de los grupos aceptores se produce más comúnmente que en grupos donadores⁶⁵. En resumen, los motivos reconocidos por la proteína están en la superficie de contacto, mientras los grupos que no interaccionan están expuestos al medio acuoso. Como los enlaces de hidrógeno de estos complejos están protegidos del disolvente, su presencia en un medio de menor constante dieléctrica determina que estas interacciones tengan una mayor contribución entálpica que los enlaces de hidrógeno con las moléculas del disolvente⁶⁶.

1.2.3. Interacciones electrostáticas

Estas interacciones están presentes en el reconocimiento de ciertos carbohidratos a través de puentes salinos entre carbohidratos cargados, como el ácido glucurónico o el ácido siálico (*NeuNAc*), y residuos de la proteína de carga opuesta. De hecho, el reconocimiento de estos azúcares cargados ha sido ampliamente abordado en aquellos receptores proteicos que interaccionan con ácido siálico, como la toxina del cólera⁶⁷, la hemaglutinina del virus de la gripe⁶⁸, o la aglutinina del germen de trigo⁵⁶. Pero estas estructuras no muestran interacciones electrostáticas entre residuos cargados positivamente y el grupo ácido del NeuNAc. En su lugar, este motivo tiene contactos con grupos amida y cadenas polares de la proteína, y moléculas de agua. En cambio, las estructuras tridimensionales de los receptores del poliomavirus⁶⁹⁻⁷⁰ o de la enzima glucuronoxilanasas⁷¹ sí muestran cómo residuos cargados positivamente, de Arg, están implicados en el reconocimiento de los grupos ácidos de los ligandos.

Un ejemplo es la neuraminidasa del virus de la gripe. Esta enzima permite la liberación de los viriones de la célula huésped, permitiendo la infección y la replicación en otras células. Ciertas drogas antivirales, como el oseltamivir (Tamiflu®) o el zanamivir (Relenza®), actúan como inhibidores en el centro ortostérico de la neuraminidasa. Sus estructuras de rayos X corroboran estas interacciones con residuos de arginina del centro activo de la enzima (Figura 11)⁷²⁻⁷⁴.

Neuraminidasa - Oseltamivir (Tamiflu®)



Neuraminidasa - Zanamivir (Relenza®)

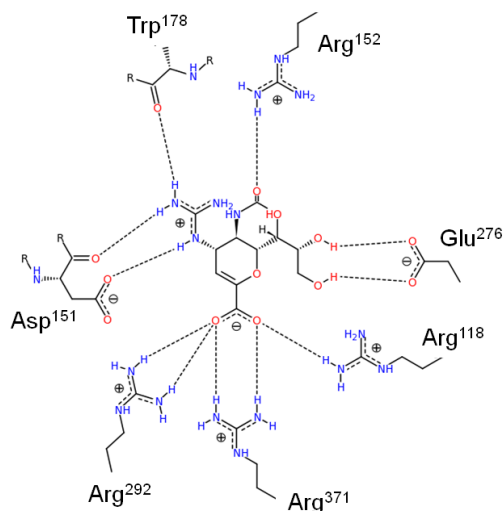


Figura 11: Representación esquemática bidimensional⁷⁵ del centro activo de la neuraminidasa de dos cepas distintas del virus de la gripe. A la *izquierda* se muestra el complejo de esta enzima con el inhibidor oseltamivir⁷⁴; a la *derecha* está representado el centro activo ocupado por una molécula de zanamivir⁷².

1.2.4. Implicación de las moléculas de agua e iones

Las moléculas de agua participan en el reconocimiento molecular a través de enlaces de hidrógeno que median entre el carbohidrato y la proteína. La comparación de estructuras de rayos X de distintos complejos de una misma proteína, o de proteínas relacionadas, muestra la posición conservada de ciertas moléculas de agua que están implicadas en la interacción. Casos estudiados en profundidad son los complejos proteína- carbohidrato de la toxina del cólera y la enterotoxina termolábil. Distintos complejos de estas proteínas con diversos ligandos han mostrado el papel fundamental de tres moléculas de agua en el proceso de reconocimiento^{67,76-79}. Asimismo, estudios de la enzima glucógeno fosforilasa han demostrado que los enlaces de hidrógeno mediados por moléculas de agua son los responsables del incremento de la afinidad cuando se une un grupo hidantoína a una molécula de glucosa⁸⁰⁻⁸¹.

Los iones también se requieren para la interacción de los carbohidratos a sus receptores. Las estructuras tridimensional de algunas lectinas han demostrado que utilizan iones de Ca^{2+} y Mn^{2+} para la estabilización de su sitio de unión, disponiendo correctamente los aminoácidos encargados del reconocimiento de sus ligandos⁸². En la *concanavalina A*, ambos metales tienen contactos con cuatro residuos de la proteína y con dos moléculas de agua. Estas uniones disponen correctamente un residuo conservado de asparagina, donador en un enlace de hidrógeno, y de un residuo de ácido aspártico, aceptor de otro.

En otras ocasiones, los iones Ca^{2+} forman directamente enlaces de coordinación con el carbohidrato, como se ha observado en las lectinas tipo C⁸³. En esta superfamilia de proteínas, dos hidroxilos vecinales del ligando forman dos enlaces de coordinación con el Ca^{2+} a través de uno de sus pares de electrónes no enlazantes del oxígeno⁸⁴⁻⁸⁶. A pesar de que el sitio de unión de estas proteínas es relativamente superficial y tiene pocos contactos con el carbohidrato, estas lectinas presentan especificidad frente al monosacárido que reconocen. Otro ejemplo en el que el ión metálico presenta contactos directos con el carbohidrato es la isomerasa de D-xilosa⁸⁷. La enzima presenta interacciones a través de uno de los iones magnesio de su centro activo con los oxígenos O2 y O4 de sus sustratos.

1.3. Glicosidasas

Este grupo de proteínas reúne una gran familia de estructuras cuya función es asistir la ruptura de un enlace glicosídico. Según su función estas proteínas son clasificadas en el sistema numérico EC (*Enzyme Commission numbers*) como EC 3.2.1.-.⁸⁸ En esta clasificación se agrupan las hidrolasas que producen la ruptura de enlaces glicosídicos de O-glucósidos y S-glucósidos⁸⁹. Este grupo es el más abundante de glicosidasas, pero existe otro grupo de enzimas que hidrolizan N-glicósidos (EC 3.2.2.-). Desde hace años, el Dr. Henrissat clasifica sistemáticamente estas enzimas mediante un análisis de similitud de secuencia de aminoácidos de las estructuras descubiertas⁹⁰⁻⁹². Las proteínas se clasifican en una misma familia si comparten un idéntico mecanismo de reacción y un plegamiento estructural semejante⁹³⁻⁹⁴.

1.3.1. Mecanismos de reacción

Excepto casos especiales⁹⁵⁻⁹⁶, estas enzimas emplean principalmente dos mecanismos de reacción para romper el enlace glicosídico. El mecanismo determina si el extremo reductor generado tras la reacción mantiene la configuración de su centro anomérico o lo invierte. Por lo tanto, una glicosidasa puede hidrolizar un enlace glicosídico mediante un *mecanismo de retención* o *de inversión*. En ambos mecanismos, el estado de transición es un ión oxocarbenio y los residuos catalíticos son dos residuos ácidos. Las glicosidasas con mecanismo de inversión tienen los dos aminoácidos catalíticos separados 10.5 Å. La hidrólisis ocurre en una reacción de un único paso en el que uno de los ácidos carboxílicos actúa como base y el otro como ácido. En cambio, las glicosidasas de mecanismo de retención tienen sus residuos separados aproximadamente 5.5 Å y la hidrólisis se produce a través de una reacción de doble desplazamiento⁹⁷.

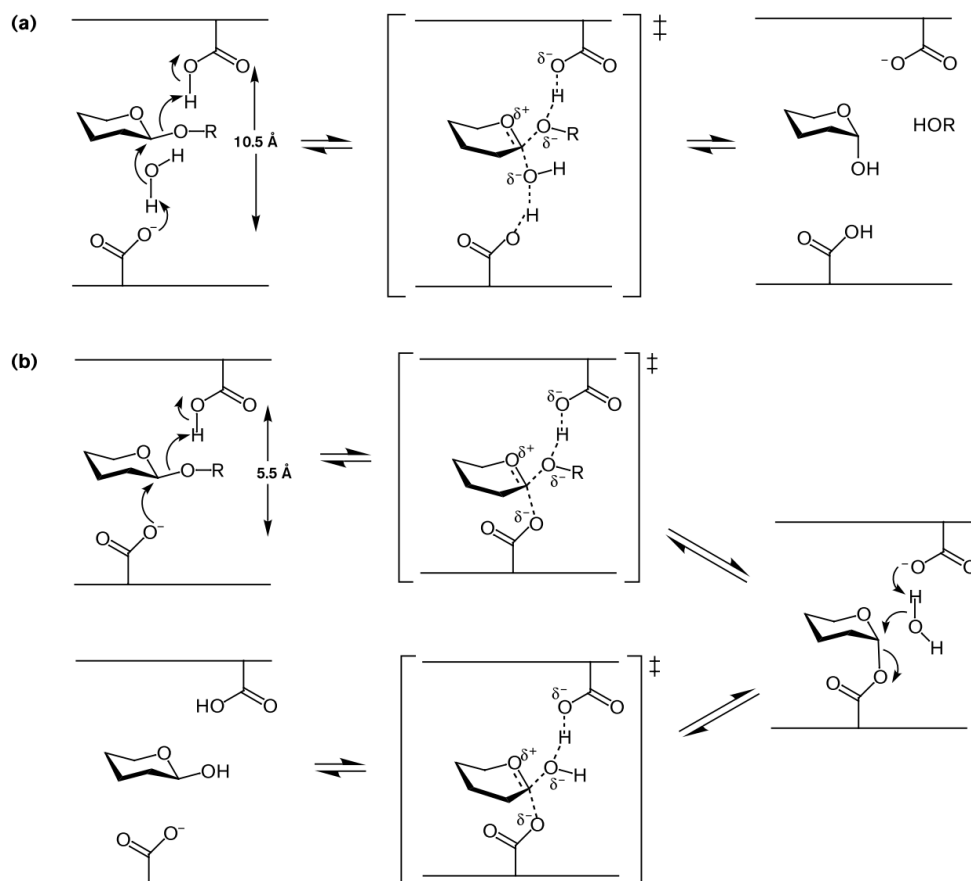


Figura 12: Esquema de los mecanismos reacción de una β -glicosidasa con mecanismo de inversión (a) y con mecanismo de retención (b). Los estados de transición de los distintas fases de reacción se representan entre corchetes (ión oxocarbenio).⁹⁷

En un *mecanismo de inversión*, uno de los residuos catalíticos actúa como ácido y el otro actúa como base. El primero asiste la hidrólisis, donando un protón, mientras el segundo activa una molécula de agua conservada que produce el ataque nucleófilo sobre el carbono anomérico del extremo reductor (Figura 12 (a)). Este proceso ocurre a través de un único estado de transición. Sin embargo, en un *mecanismo de retención* se observa un estado intermedio durante la reacción enzimática. Este otro mecanismo se divide en dos fases, fase de glicosilación y fase de desglicosilación. En la fase de glicosilación, el ataque nucleófilo lo realiza un residuo de la enzima, permaneciendo el extremo reductor del enlace glicosídico anclado a la proteína. La siguiente etapa, fase de desglicosilación, produce la liberación del producto. La proteína glicosilada es hidrolizada por una molécula de agua, activada por el grupo básico catalítico que actuaba como ácido en la fase de glicosilación (Figura 12 (b))⁹⁸.

1.3.2. Inhibidores de glicosidasas

El modo de acción de los inhibidores de estas enzimas está íntimamente relacionado con su mecanismo de acción. En una primera clasificación, los inhibidores ortostéricos de las glicosidasas se diferencian en irreversibles y reversibles. La interacción de los inhibidores irreversibles genera complejos covalentes tras la unión, bloqueando el centro activo de la enzima de manera permanente. En cambio, los reversibles se reconocen por la hidrolasa sin reaccionar con ningún residuo del centro activo.

Los inhibidores irreversibles clásicos sufren el ataque nucleófilo de un residuo de la hidrolasa. El complejo generado impide el acceso del sustrato al sitio activo de la enzima. Estas moléculas han sido utilizadas para la identificación de los residuos catalíticos y su diferenciación de aquellos cuya función es estructural. Además, se han utilizado para inactivar distintas dianas enzimáticas⁹⁹. Por ejemplo, los 2-desoxi-2-fluoro glicósidos (Figura 13-1) actúan desestabilizando el estado de transición oxocarbenio. Estos glicósidos fluorados contienen grupos salientes muy efectivos que favorecen la etapa de glicosilación. Además, la desglicosilación está muy desfavorecida, por lo que se acumula el estado intermedio¹⁰⁰⁻¹⁰¹. Estos inhibidores son una herramienta precisa para la identificación del residuo nucleófilo, y ha permitido identificar residuos catalíticos poco habituales, como Thr¹⁰² o Tyr¹⁰³. Otros inhibidores irreversibles, como los glicósidos N-bromo-acetilados (Figura 13-2), son capaces de etiquetar el residuo catalítico ácido/base de ciertas β -glicosidasas¹⁰⁴⁻¹⁰⁶. Otros inhibidores covalentes son los epóxidos y la aziridinas. Un miembro de esta familia es el epóxido del conduritól B (Figura 13-3), inhibidor de la glucocerebrosidasa humana que actúa en la degradación de los glicoesfingolípidos. Este inhibidor selectivo¹⁰⁷ es una herramienta potente para el diseño de fármacos frente a trastornos metabólicos como la enfermedad de Gaucher¹⁰⁸⁻¹⁰⁹.

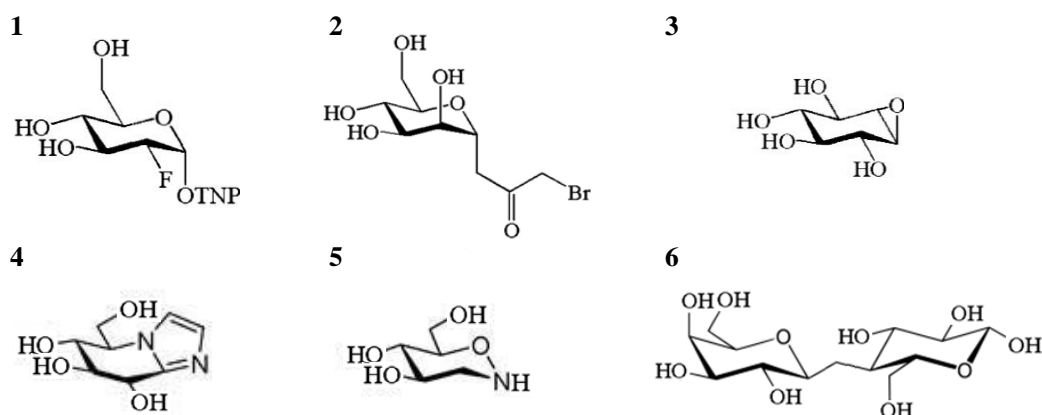


Figura 13: Estructura de distintos inhibidores de glicosidasas: trinitrofenil 2-desoxi-2-fluoro glucósido (1), bromoacetona C-manósido (2), conduritól B-epóxido (3), glucoimidazol (4), tetrahidrooxazina (5), y C-lactosa (6).

Los inhibidores reversibles no se unen covalentemente, pero sus constantes de inhibición pueden alcanzar valores de alta afinidad¹¹⁰. Según su estructura, estos

inhibidores se pueden clasificar en dos clases, análogos del estado de transición y análogos de sustrato. El grupo de análogos del estado de transición está formado por diversas estructuras que mimetizan la estructura trigonal del ión oxocarbenio. Ejemplos de estos inhibidores son la gluconolactona, la nojirimicina, la castanospermina, la isofagomina, el glucotetrazol, el glucoimidazol (Figura 13-4), la acarbosa¹¹¹. En principio, una molécula que mimetiza el estado de transición debe tener una interacción dirigida entálpicamente¹¹²⁻¹¹³. Este análisis, propuesto por Wolfenden, se ha aplicado a un sistema modelo, la β -glucosidasa de *Thermotoga maritima*. Así, se ha identificado la tetrahidrooxazina (Figura 13-5) como el compuesto más representativo de los análogos del estado de transición analizados¹¹⁴. El análisis teórico de la estabilización producida por el estado de transición demuestra que la unión de un inhibidor análogo podría alcanzar un constante de disociación de 10^{-20} M. Sin embargo, todavía no se ha diseñado un inhibidor capaz de mimetizar con tanta precisión la forma y la carga del estado de transición¹¹⁵.

Dentro del otro grupo de inhibidores reversibles, los análogos de sustratos, hay que destacar los C-glicósidos y los tioglicósidos (S-glicósidos). La sustitución del oxígeno interglicosídico por un átomo de carbono o azufre previene la hidrólisis enzimática, y mantiene su capacidad de interacción sin ser procesados. El reconocimiento molecular de estos inhibidores puede modificar la conformación del anillo de su extremo reductor o la disposición de sus ángulos glicosídicos. Por ejemplo, la conformación mayoritaria de la C-lactosa (Figura 13-6) en disolución no corresponde con la disposición reconocida por la β -galactosidasa de *E.coli*. La C-lactosa adopta una conformación *anti* en su torsional ϕ cuando interacciona, a pesar de ser una conformación minoritaria en estado libre¹¹⁶⁻¹¹⁷. En cambio, la β -glucosidasa de *Streptomyces* sp. reconoce el conformero mayoritario de la tiocelobiosa en estado libre. Tanto el inhibidor como el sustrato natural, la celobiosa, adoptan una conformación *syn* en el centro activo. Por lo tanto, ambos compuestos se acomodan del mismo modo en el centro activo¹¹⁸. Obviamente, esta selección conformacional depende tanto de la estructura del inhibidor como de la enzima.

1.4. Resonancia Magnética Nuclear (RMN)

La elucidación estructural de moléculas biológicas se ha realizado principalmente mediante dos técnicas, la difracción de Rayos X y la Resonancia Magnética Nuclear. Estas técnicas presentan notables diferencias en la distribución de las moléculas en la muestra que se observa y en las escalas de tiempo que son accesibles para cada una de ellas. La RMN en disolución obtiene información de un promedio de moléculas orientadas en un disolvente, mientras que la difracción de rayos X utiliza el promedio de la disposición de las moléculas en una red cristalina periódica. El empaquetamiento de esta red cristalina puede afectar a la estructura, fijándola en una conformación particular. Sin embargo, en principio, los estudios en disolución no afectan a la posibilidad de que la molécula adopte diferentes conformaciones¹¹⁹. En cuanto a las

escalas de tiempo, los datos de RMN cubren un intervalo desde los picosegundos a los segundos. En cambio, los datos de difracción cubren una escala de tiempo desde los segundos a las horas. El tamaño molecular de los sistemas que pueden ser estudiados por cristalografía de rayos X cubre un espectro mucho más amplio, desde moléculas de bajo peso molecular hasta complejos macromoleculares, como el ribosoma. Sin embargo, los estudios de RMN tienen un límite en el tamaño molecular accesible. A pesar de las actuales mejoras experimentales, no se han elucidado todavía estructuras de un peso molecular superior a los 100 kDa, aunque sí ha sido posible registrar espectros de supramoléculas de este tamaño. A pesar de este límite experimental, la RMN permite el análisis de estructuras de alta flexibilidad o la caracterización experimental de regiones de mayor flexibilidad dentro de una macromolécula. Además, la RMN permite el estudio de detalles del plegamiento de proteínas mediante experimentos de intercambio químico. Estas técnicas son complementarias, la combinación de RMN y difracción de Rayos X permite obtener modelos estructurales más precisos. En esta tesis doctoral se han usado diversas técnicas de RMN, debido a que el tamaño de las moléculas es adecuado y los sistemas analizados son altamente flexibles¹²⁰⁻¹²¹.

1.4.1. Aspectos generales de RMN

La técnica experimental que principalmente se ha utilizado en esta tesis doctoral es la *Resonancia Magnética Nuclear*, RMN¹²²⁻¹²³. El instrumento es un imán superconductor capaz de generar un potente campo magnético que alinea los núcleos magnéticos haciéndolos girar con un movimiento de precesión, *precesión de Larmor*. Los núcleos con momento magnético más utilizados en RMN son ^1H , ^{13}C , ^{15}N , ^{19}F o ^{31}P . Los espines de un tipo de núcleo se comportan en la muestra de manera similar, pero hay diferencias según sea la posición del átomo dentro de la molécula. Estas pequeñas variaciones dependientes del ambiente que rodea al núcleo determinan que cada uno de estos núcleos tenga un *desplazamiento químico* característico.

Los núcleos habitualmente utilizados en RMN, como el ^1H , tienen dos estados cuánticos, $+1/2$ y $-1/2$ ó α y β . Además, los espines están sujetos a los efectos debidos a la proximidad de otros núcleos magnéticamente activos. Cada núcleo se comporta como un pequeño imán que afecta a su vecino por un fenómeno denominado acoplamiento de espines (*spin-spin splitting* o *spin-spin coupling* en inglés). Considerando un sistema con dos núcleos adyacentes que poseen dos estados cuánticos, el campo magnético al que están sometidos se perturba de distinta manera según si el núcleo contiguo tiene un espín α o β . Consecuentemente, las señales del espectro de RMN adoptan una apariencia de *doblete*. La distancia entre los dos picos del *doblete* se denomina *constante de acoplamiento*, J . Este mismo fenómeno que hemos explicado para un sistema sencillo de dos núcleos acoplados puede extenderse a sistemas complejos (Figura 14).

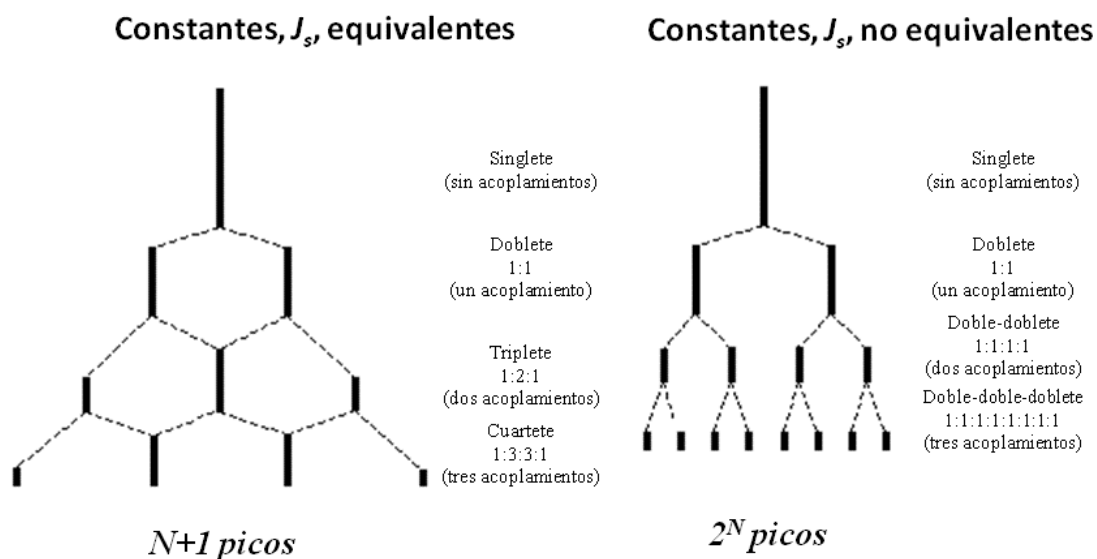


Figura 14: Representación en árbol de distintos patrones de acoplamiento que pueden observarse para una señal de resonancia dependiendo de la cantidad de acoplamientos y de la equivalencia o no de sus valores.

La separación entre los distintos picos del multiplete son las *constantes de acoplamiento* que se miden en Hertzios, Hz. Las constantes de acoplamiento escalar dan información estructural de la molécula ya que su magnitud depende de la disposición relativa de los dos núcleos. Las constantes de acoplamiento entre dos protones situados a tres enlaces, $^3J_{HH}$, están relacionadas con su ángulo diedro mediante una ecuación de tipo Karplus (El valor de la constante también depende de las electronegatividades de los distintos sustituyentes enlazados. Por eso, los valores A, B y C varían según las sustituciones que se presenten¹²⁴).

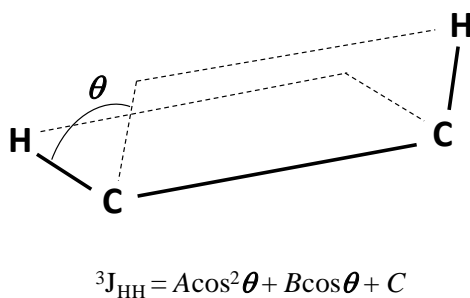


Figura 15: Dependencia estereoquímica de la constante de acoplamiento a tres enlaces expresada mediante la ecuación propuesta por Karplus. Esta ecuación permite calcular la constante de acoplamiento esperada entre dos protones cuyo ángulo diedro tiene el valor θ . Los valores A, B y C son constantes ajustadas empíricamente.

Esta dependencia estereoquímica de las constantes de acoplamiento se utiliza tanto en la identificación de distintas configuraciones como en análisis conformacional. Por ejemplo, la geometría de los monosacáridos puede ser determinada por el análisis de sus constantes de acoplamiento¹²⁵. Cuando ambos protones están en disposición axial su

constante de acoplamiento tiene un valor de 7-9 Hz. En cambio, si al menos uno de ellos tiene una disposición ecuatorial este valor es significativamente menor. Además, este parámetro experimental puede ser utilizado para el estudio de equilibrios conformacionales, ya que las constantes de acoplamiento observadas dependerán del peso poblacional de cada uno de los conformeros.

La complejidad de un espectro monodimensional de RMN en las moléculas biológicas determina que sea necesaria la utilización de experimentos bidimensionales de protón y carbono que permitan la asignación de cada señal a su correspondiente átomo. Hay espectros de correlación escalar que permiten la correlación entre espines a través de los enlaces químicos¹²⁶⁻¹²⁷. Los dos experimentos homonucleares de este tipo más comunes son los calificados *Correlation Spectroscopy* o COSY, y *Total Correlation Spectroscopy* o TOCSY. El experimento COSY correlaciona aquellos protones que están acoplados a dos y tres enlaces. Por lo tanto, es posible crear una red de conexiones entre los espines que están consecutivamente acoplados pudiendo identificar los distintos *sistemas de espín* de la molécula. Un *sistema de espín* está constituido por el grupo de espines que interactúan entre ellos en un campo magnético a través de constantes de acoplamiento. Los sistemas de espín de una molécula también se identifican mediante experimentos TOCSY que relacionan cada espín con el resto de espines de un mismo sistema. Además de estos experimentos homonucleares de correlación escalar, también se emplean experimentos heteronucleares que correlacionan distintos núcleos a través de sus enlaces químicos, como el HSQC y el HMBC, o experimentos homonucleares que correlación espines a través del espacio, como el NOESY o el ROESY.

1.4.2. Relajación dipolo-dipolo y efecto Overhauser nuclear

Este mecanismo de relajación entre espines proporciona la información espacial, que es monitorizada en ciertos experimentos de RMN. La comprensión de este mecanismo puede ser abordada con la ayuda de un sistema simplificado, que ha sido ampliamente descrito y se representa en la Figura 16.¹²⁸⁻¹²⁹

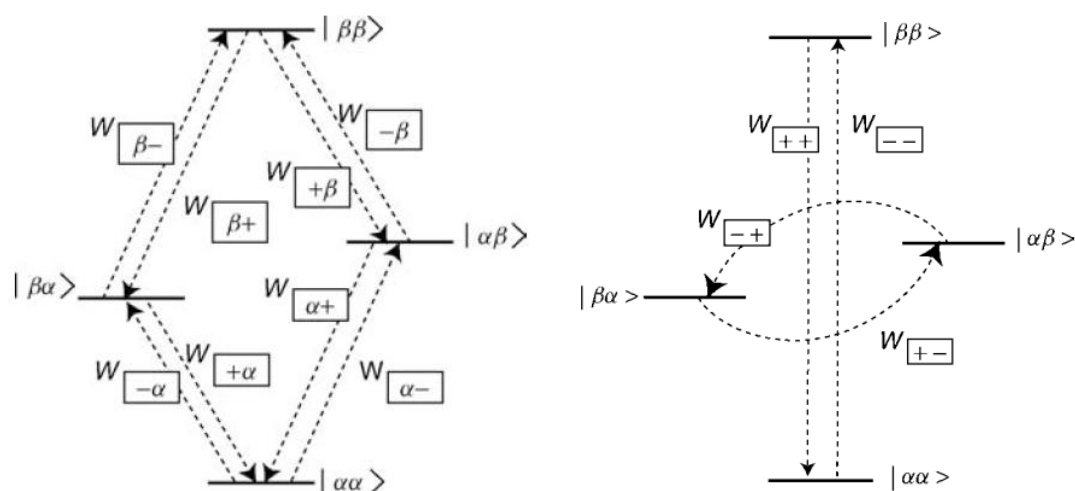


Figura 16: *Izda.*, las posibles transiciones de *single-quantum* en un sistema homonuclear. *Dcha.*, las posibles transiciones de *doble-quantum* (W_{++} y W_{--}) y *zero-quantum* (W_{+-} y W_{-+}) en un sistema homonuclear.

El efecto de la relajación dipolo-dipolo se traduce en el denominado *efecto Overhauser nuclear (NOE)*. La probabilidad de estas transiciones energéticas es inversamente proporcional a la sexta potencia de la distancia (distancias menores de 5 Å). Esta dependencia de la distancia internuclear es utilizada en la determinación estructural de moléculas en RMN¹²⁸⁻¹²⁹. Además, si la constante de relajación cruzada, σ , es positiva el signo del pico de cruce será opuesto al del pico diagonal en los experimentos NOESY. Si σ es negativa, será del mismo signo. El signo de σ depende del tiempo de correlación de la molécula, τ_c , y de la frecuencia de Larmor, ω .

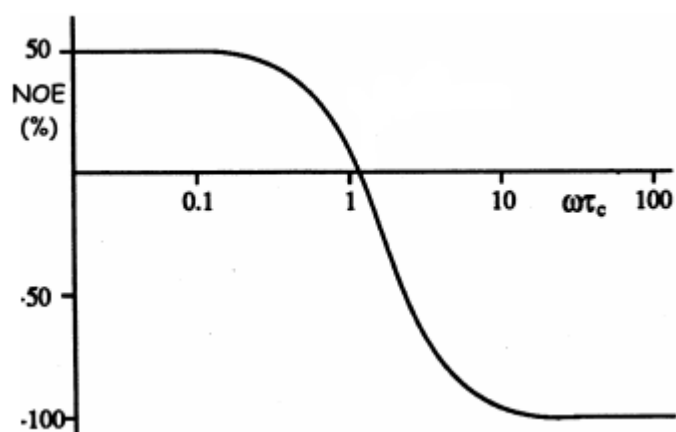


Figura 17: Dependencia del NOE frente al producto $\omega\tau_c$. ω es la frecuencia del Larmor y τ_c el tiempo de correlación de la molécula. Se representa el porcentaje de NOE, y su signo viene referenciado a una señal diagonal de signo negativo.

El tiempo de correlación, τ_c , se relaciona directamente con el tamaño de la molécula, las moléculas pequeñas tienen τ_c cortos y las grandes, τ_c largos. Si $\omega\tau_c < 1$, condición de molécula pequeña (MW < 2000 Da), los picos de cruce son de signo opuesto a la

diagonal y si $\omega\tau_c > 1$, condición de molécula grande, los picos de cruce son del mismo signo.

Por otro lado, la relajación cruzada entre dos espines también puede medirse mediante experimentos ROESY (*Rotating-frame Overhauser Effect Spectroscopy*). En estos experimentos, la relajación cruzada ocurre en el plano transversal bajo la influencia del campo B_1 , mientras que en los NOESY, tiene lugar en el eje z, bajo la influencia del campo estático, B_0 . El campo generado por el tren de pulsos, B_1 , es de menor magnitud que el estático, B_0 . Por lo tanto, $\omega_1\tau_c \ll 1$ para cualquier tamaño molecular. Por lo tanto, los picos de cruce de ROESY son siempre de signo opuesto a la diagonal (ROE directo) para cualquier tiempo de correlación, τ_c (Figura 18).

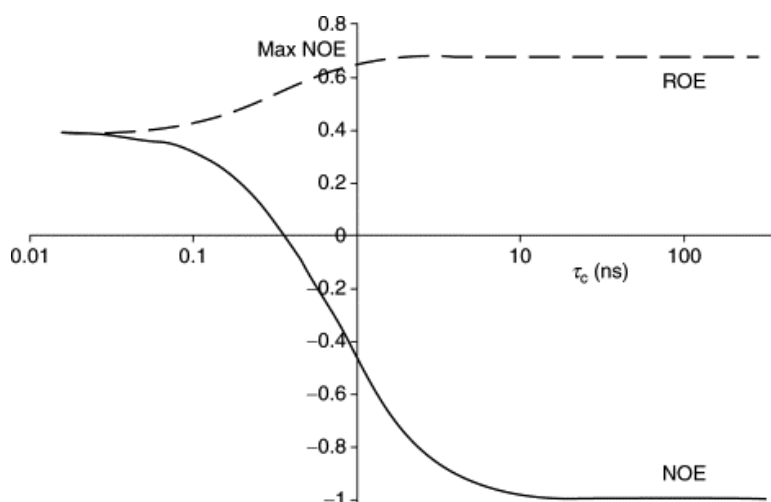


Figura 18: Máximos valores obtenidos en los experimentos NOESY (línea continua) y ROESY (línea discontinua) en función del tiempo de correlación, τ_c . El efecto NOE es 0,38 para moléculas pequeñas y -1.0 para moléculas grandes. Sin embargo, el efecto ROE alcanza un valor máximo de 0,68 para moléculas grandes.

origen del pico cruce	Signo del pico de cruce - NOESY	Signo del pico de cruce - ROESY
<i>NOE/ROE directo</i>	Positivo (\downarrow MW) y Negativo (\uparrow MW)	Positivo
<i>NOE/ROE indirecto</i>	Positivo (\downarrow MW) y Negativo (\uparrow MW)	Negativo
<i>TOCSY</i>	-----	Negativo
<i>TOCSY-ROE</i>	-----	Positivo (falso ROE)
<i>ROE-TOCSY</i>	-----	Positivo (falso ROE)
<i>intercambio químico</i>	Negativo	Negativo
<i>tipo COSY</i>	Pico en antifase	Pico en antifase

Tabla 1: Origen y signo de los picos de cruce que pueden encontrarse en los experimentos NOESY y ROESY. El signo del pico de cruce está referenciado a un pico diagonal de signo negativo.

1.4.3. Asignación de péptidos por RMN

La estructura lineal de los péptidos y la separación de sus aminoácidos por los enlaces amida determina que cada uno de los aminoácidos se comporte como un sistema de espín en RMN. Mediante experimentos bidimensionales de correlación escalar, COSY y TOCSY, se asignarán los distintos sistemas de espín y, seguidamente, la secuencia de aminoácidos será trazada por espectros de correlación espacial, NOESY y ROESY.

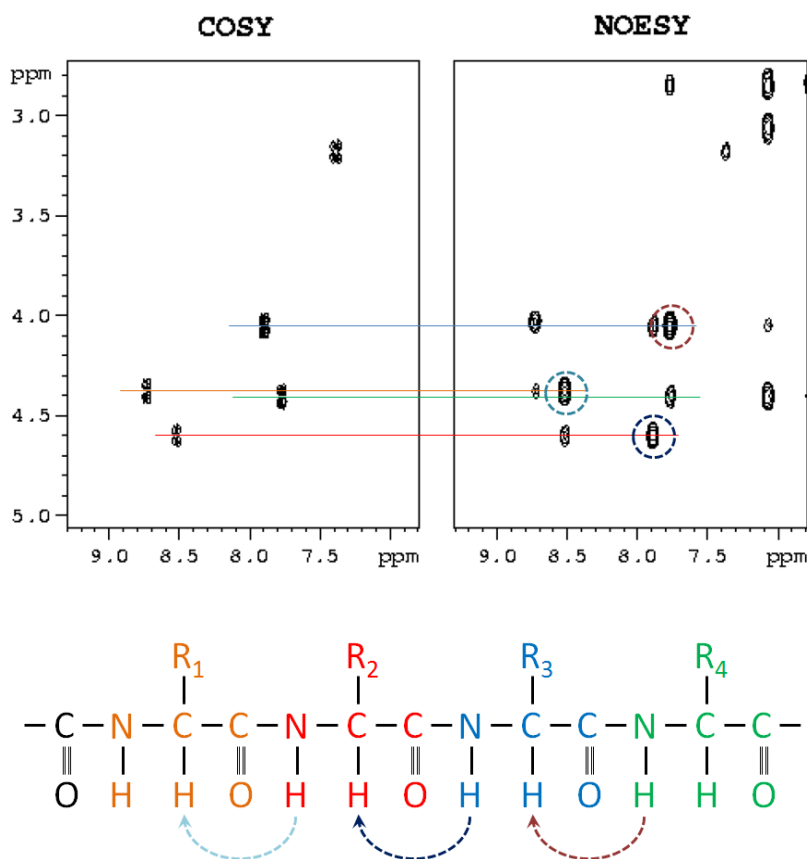


Figura 19: *Arriba*, espectros COSY y NOESY de un péptido modelo. La conectividad puede ser trazada por los picos de cruce entre el grupo amida de un aminoácido y el protón alfa del anterior. *Abajo*, representación lineal de la cadena de aminoácidos del péptido (NOEs interresiduales del esqueleto peptídico se representan con flechas).

Los aminoácidos naturales de un péptido pueden identificarse según su patrón de señales. En la Figura 20 se representa, de manera esquemática, la información escalar que se obtiene de los experimentos COSY y TOCSY de los 20 aminoácidos naturales. En estos dos experimentos el protón amídico de un aminoácido sólo tiene correlación con uno de los protones α del péptido, sólo está acoplado escalarmente con su propio protón α . Sin embargo, en los experimentos NOESY y ROESY, éste mismo protón amídico muestra la correlación con su propio protón α y el del anterior residuo en la secuencia. La comparación de estos dos tipos de experimentos permite el trazado de la secuencia de aminoácidos (Figura 19)¹³⁰.

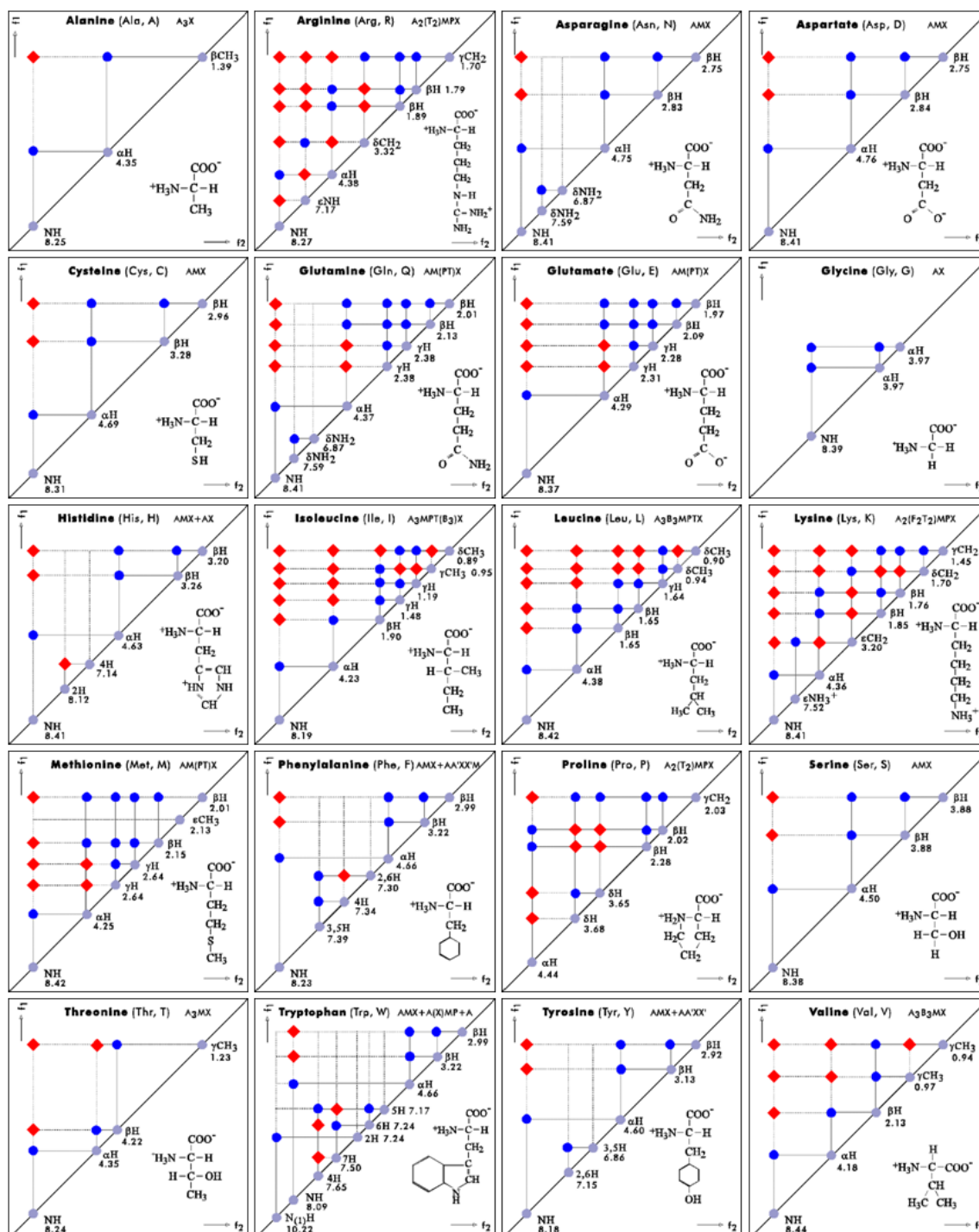


Figura 20: Representación de los experimentos COSY y TOCSY de los 20 aminoácidos naturales. Los picos de cruce COSY son representados con círculos azules y los picos de cruce TOCSY están simbolizados con rombos rojos.

Como se ha mencionado en el apartado de la relajación dipolo-dipolo, la magnitud del pico de cruce de un espectro NOESY es inversamente proporcional a la sexta potencia

de la distancia entre los núcleos. La utilización de unidades arbitrarias en el espectro determina que sea necesaria una referencia interna de distancia. Esta referencia puede ser una pareja de protones geminales, una pareja de dos protones vecinos aromáticos o una pareja de protones 1,3-diaxial. En algunas ocasiones se pueden encontrar picos de cruce entre protones que están más lejanos de la distancia observable en un espectro NOESY, mayores de 5 Å. Esta observación experimental viene determinada por un proceso de NOE indirecto denominado *difusión de espín*. Consiste en la perturbación de las poblaciones de los espines vecinos (NOE indirecto) tras una previa perturbación de la magnetización de un protón (NOE directo). Este proceso aumenta su probabilidad cuando se usan tiempos de mezcla elevados. Por lo tanto, es recomendable la utilización de tiempos de mezcla cortos para disminuir la *difusión de espín*. Si se generan curvas de NOE a tiempos de mezcla crecientes, se puede discernir fácilmente un NOE directo de un NOE indirecto, ya que este último muestra una curva sigmoidea.

Para el cálculo de distancias se aplica la fórmula siguiente:

$$r_x = r_{ref} \left[\frac{\sigma_{ref}}{\sigma_x} \right]^{1/6}$$

σ_{ref} es la integral del pico de cruce entre los dos espines a una distancia referencia r_{ref} , σ_x es la integral del pico de cruce entre los dos espines cuya distancia queremos medir y r_x es la distancia a calcular.

La información de los picos de cruce obtenida de un espectro NOESY es difícilmente procesada sin la ayuda de programas que faciliten la transformación de estos datos experimentales en información estructural. Por ejemplo, en el análisis conformacional de péptidos se utiliza el programa CYANA¹³¹. Este programa permite el cálculo estructural a partir de los picos de cruce observados en los espectros NOESY de un péptido, que previamente ha sido asignado usando el software CARA¹³². En otras ocasiones no es necesario un cálculo cuantitativo. Usando un tiempo de mezcla del orden de los valores de la relajación longitudinal, T1, se optimiza la intensidad de los picos de cruce NOE. Sin embargo, en estas condiciones, la intensidad no es estrictamente proporcional a $1/r^6$, y debe considerarse la difusión de espín. Con esta aproximación, los picos de cruce NOE se clasifican en *débiles*, *medios* y *fuertes*, y se les asigna unos rangos de distancia aproximados. Esta aproximación se utiliza muy frecuentemente para el cálculo de la estructura de proteínas.

1.5. RMN y reconocimiento molecular

La comprensión del proceso de reconocimiento entre los ligandos y sus dianas es clave para el diseño de inhibidores. La RMN permite identificar qué compuestos interaccionan con la macromolécula, qué conformación adoptan en el centro de unión y qué regiones están implicados en la unión. Los métodos utilizados se dividen en dos

categorías: experimentos desde el punto de vista del receptor o desde la perspectiva del ligando¹³³⁻¹³⁵. Los métodos basados en la observación del ligando se basan en las distintas propiedades que tienen estos compuestos de bajo peso molecular en estado libre y unido a sus receptores. En estado libre, tienen velocidades de relajación, $R1$ y $R2$, lentas, los picos de cruce de los experimentos NOESY presentan signo opuesto a los picos de la diagonal, y sus coeficientes de difusión translacional son altos. En cambio, estos compuestos al asociarse a una macromolécula adoptan sus propiedades. En estado unido, el ligando muestra relajación rápida, los picos de cruce NOESY son del mismo signo que los picos de la diagonal, y el coeficiente de difusión es bajo (Figura 21).

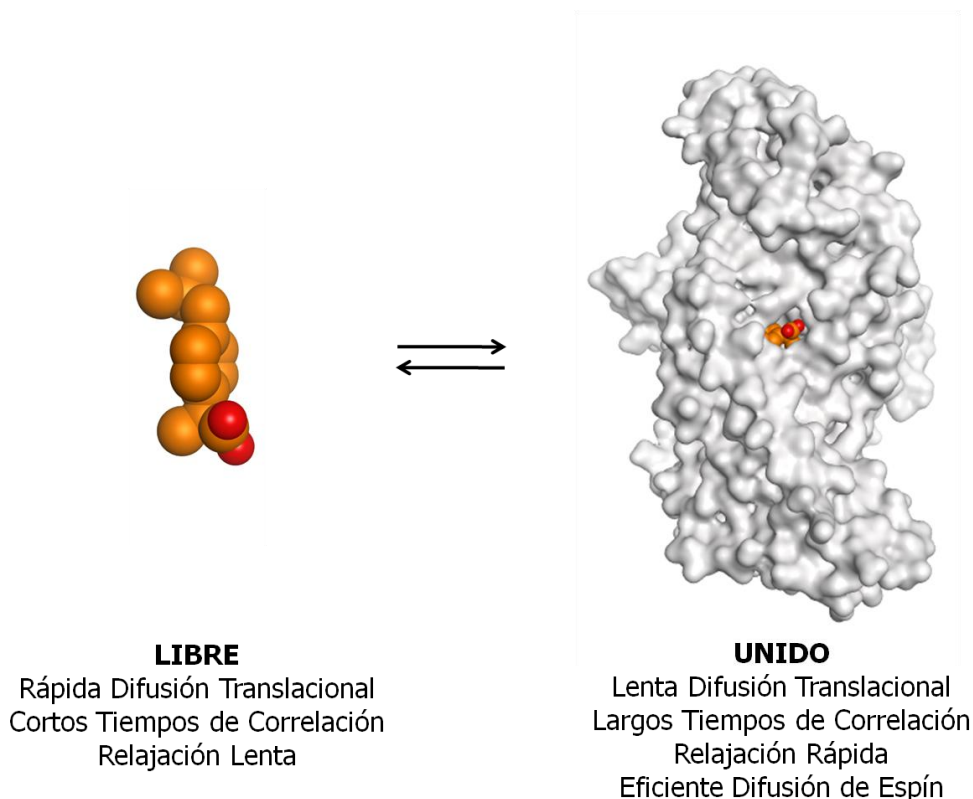


Figura 21: Equilibrio entre el estado libre del ibuprofeno y la forma unida a su receptor, albúmina de suero bovino, *BSA*. La molécula de ibuprofeno en estado libre tiene propiedades de molécula pequeña, que se modifican si su receptor está presente en la muestra. Cuando un ligando es reconocido, difunde lentamente, sus NOE son del mismo signo que la diagonal (negativos) y su relajación es rápida, como en la macromolécula.

En esta tesis doctoral se han utilizado experimentos que analizan el reconocimiento molecular mediante técnicas de RMN que se centran en la observación de las señales del ligando. Se han empleado principalmente espectros de diferencia de transferencia de saturación (STD), espectros bidimensionales de NOE transferido (TR-NOESY) y experimentos de difusión (DOSY). Estos métodos pueden dividirse en dos grupos, según: los ligados a un cambio en los movimientos rotacionales o translacionales.

1.5.1. Métodos basados en cambios en los movimientos rotacionales

1.5.1.1. Espectroscopia de diferencia de saturación (STD)

El STD (*Saturation Transfer Difference*)¹³⁶ es el experimento de RMN más usado para identificar la unión de un ligando a un receptor. Permite determinar si hay asociación y cuáles son las regiones del ligando que se encuentran en contacto directo con la proteína en el complejo (epítipo). La técnica se basa en la saturación de las resonancias pertenecientes a la proteína receptora mediante un tren de pulsos selectivos. Mediante difusión de spin, esta saturación se transfiere al resto de protones de la proteína, incluidos los protones de los residuos del sitio de reconocimiento, que terminan afectando a las resonancias del ligando unido. Cuando el equilibrio entre el estado libre y el estado asociado es rápido en la escala de tiempos de relajación, la transferencia de la magnetización que se produce en el ligando asociado se transmite al ligando libre, produciendo una disminución en la intensidad de sus resonancias. El efecto observado depende del tamaño de la molécula, de la constante de asociación, de la constante de velocidad de disociación y del tiempo de saturación.

Un experimento STD es la diferencia entre dos espectros monodimensionales: el espectro con irradiación selectiva de las resonancias de la proteína, *on-resonance*, y el espectro cuya saturación se ha realizado en una región sin señales de la proteína ni del ligando, *off-resonance*. La diferencia entre los experimentos *on-resonance* y los *off-resonance* es el espectro STD. Este muestra únicamente las señales de los protones más cercanos al receptor en su conformación unida y, por lo tanto, el epítipo (Figura 22). Las concentraciones empleadas de receptor son bajas. Además, puesto que la relajación es rápida, sus señales no suelen ser visibles en el STD. Si lo son, pueden ser eliminadas fácilmente empleando un filtro de relajación antes de la detección.

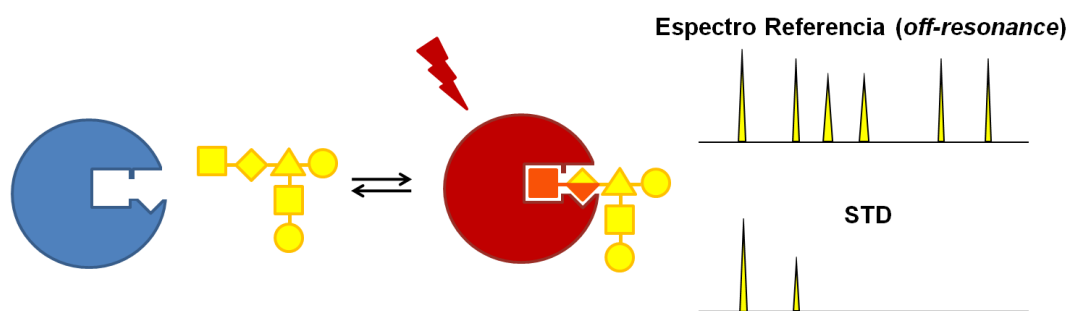


Figura 22: Representación de un experimento STD. La saturación en la proteína se transfiere al ligando, si hay interacción, mediante difusión de espín. El efecto STD es más acusado en las regiones del ligando más íntimamente unidas a la macromolécula.

1.5.1.2. Secuencias WaterLOGSY

WaterLOGSY (*Water-Ligand Observed via Gradient Spectroscopy*)¹³⁷ también utiliza la transferencia de magnetización tras una irradiación selectiva. En esta ocasión se invierten las señales de las moléculas de agua de la muestra. Durante el tiempo de mezcla se produce la transferencia al ligando mediante relajación dipolo-dipolo. La magnetización es recibida directamente desde las moléculas de agua que ocupan la superficie de contacto entre el ligando y la macromolécula, o indirectamente desde aquellos protones lábiles de la proteína que están en intercambio químico con los protones del agua. Durante el tiempo de mezcla, tanto los protones intercambiables de la proteína como los protones de las moléculas de agua que forman parte del complejo ligando-proteína adquieren las propiedades de la macromolécula. Por lo tanto, como ocurre en un experimento NOESY, el aumento del tiempo de correlación de estos protones determina que el signo de la señales de los ligandos unidos sea opuesto al del resto de compuestos de una mezcla (). Los compuestos que no son reconocidos por el receptor reciben la magnetización de las moléculas de agua de su esfera de solvatación, que muestran un tiempo de correlación muy rápido¹³⁸.

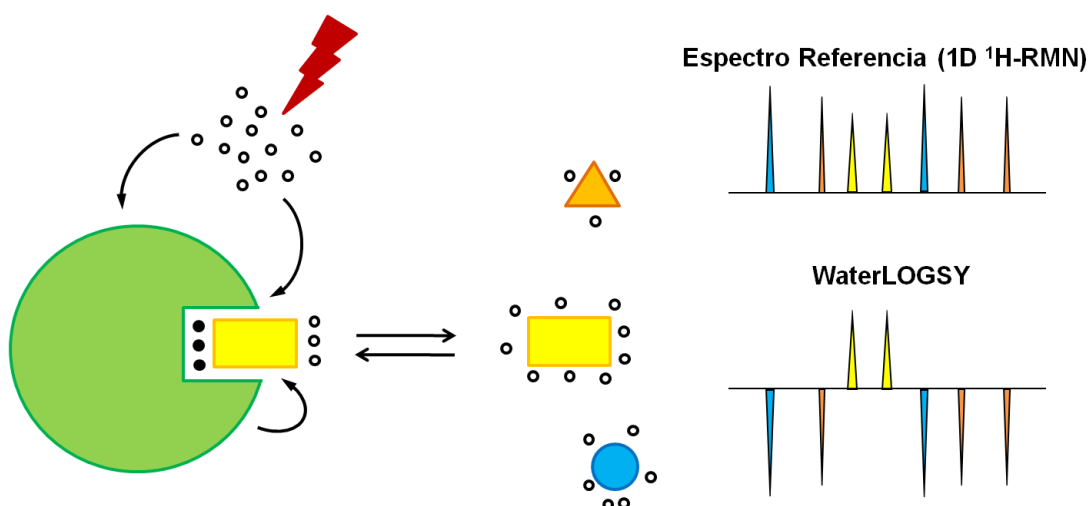


Figura 23: Esquema del experimento WaterLOGSY. Las moléculas de agua de la esfera de solvatación se representan con círculos huecos, y la moléculas de la superficie de contacto entre el ligando y la macromolécula con círculos negros. A la derecha se muestra la comparación entre un espectro de protón y un experimento WaterLOGSY. En el WaterLOGSY, las señales del ligando presentan signo contrario a las de las señales del resto de compuestos que no interaccionan con la proteína¹³⁹.

1.5.1.3. NOE-transferido (TR-NOE).

Normalmente, los ligandos son moléculas de bajo peso molecular. Estas moléculas tienen un comportamiento hidrodinámico caracterizado por tiempos de correlación cortos, que determinan velocidades de relajación lentas. Así, sus picos de cruce en experimentos NOESY son positivos, de distinto signo que los picos de la diagonal. En cambio, las moléculas de alto peso molecular tienen elevados tiempos de correlación y,

por tanto, velocidades de relajación rápidas. En este caso, los picos de cruce NOESY tienen el mismo signo que los picos de la diagonal (Figura 24)¹⁴⁰.

Cuando un ligando interacciona con una macromolécula, cambia su comportamiento y adopta, al menos durante una fracción de tiempo, las propiedades de la molécula de alto peso molecular. Lógicamente, el tiempo de residencia en el centro de reconocimiento dependerá de la constante de velocidad de disociación, k_{off} . Por tanto, el ligando unido a su receptor presentará velocidades de relajación rápidas y picos de cruce NOESY negativos. La comparación entre los espectros en estado libre y unido permite obtener información estructural de las conformaciones en ambos estados y descifrar posibles cambios conformacionales que pueden ocurrir durante el proceso de reconocimiento.

Para un par de núcleos, I y S, de un ligando cuyo proceso de intercambio entre la forma libre (L) y la unida (U) es rápido en la escala de desplazamientos químicos, el NOE observado es un promedio de la contribución de los NOEs de ambos estados, que se expresa en función de la constante de relajación promedio (σ_{IS}):

$$\langle \sigma_{IS} \rangle = N^L \sigma_{IS}^L + N^U \sigma_{IS}^U$$

σ_{IS}^L y σ_{IS}^U son las constantes de relajación cruzada en estado libre y unido; N^L y N^U son las fracciones molares respectivas. Por tanto, si en esta expresión domina el término de la forma unida, los NOEs observados caracterizarán a la conformación del ligando en el estado unido, mientras que si el término que domina es el del estado libre, los NOEs reflejarán la conformación del ligando en dicho estado. Por lo tanto, el experimento TR-NOESY contiene la información deseada siempre que se cumpla:

$$|N^f \sigma_{IS}^f| \ll |N^b \sigma_{IS}^b|$$

La constante de relajación cruzada (σ) está relacionada con el tiempo de correlación (τ_c). Aquellos compuestos cuyo peso molecular es medio-bajo tienen un τ_c muy rápido (decenas de ps) y, como consecuencia, exhiben NOEs positivos. Al interactuar, la molécula pequeña se une a su diana y su tiempo de correlación efectivo varía (ns o incluso μ s). La rotación lenta determina que los picos de cruce NOE pasen a ser negativos. Las condiciones para que el TR-NOE se produzca serán más favorables cuanto más grande sea el receptor, ya que un mayor tamaño del complejo determina un tiempo de correlación mayor. Asimismo, es necesario que la velocidad de disociación del complejo sea rápida en la escala de tiempos de la relajación. De ese modo, se produce la transferencia de información entre los dos estados. El rango óptimo de afinidades para estos experimentos está entre 100 nM y 10 mM.

La difusión de espín puede afectar a la información que proviene de un experimento TR-NOESY. En algunas ocasiones, además del intercambio directo de magnetización entre protones cercanos en el espacio, se produce difusión de spin mediada por protones

del receptor o del propio ligando. Este hecho puede conducir a una interpretación incorrecta de los espectros, ya que los picos de cruce entre estos protones lejanos también tendrían el mismo signo que los picos de la diagonal en un TR-NOESY. La utilización de TR-ROESY permite distinguir entre aquellos picos de cruce que provienen de un contacto directo y los dominados por efectos indirectos (Figura 24).

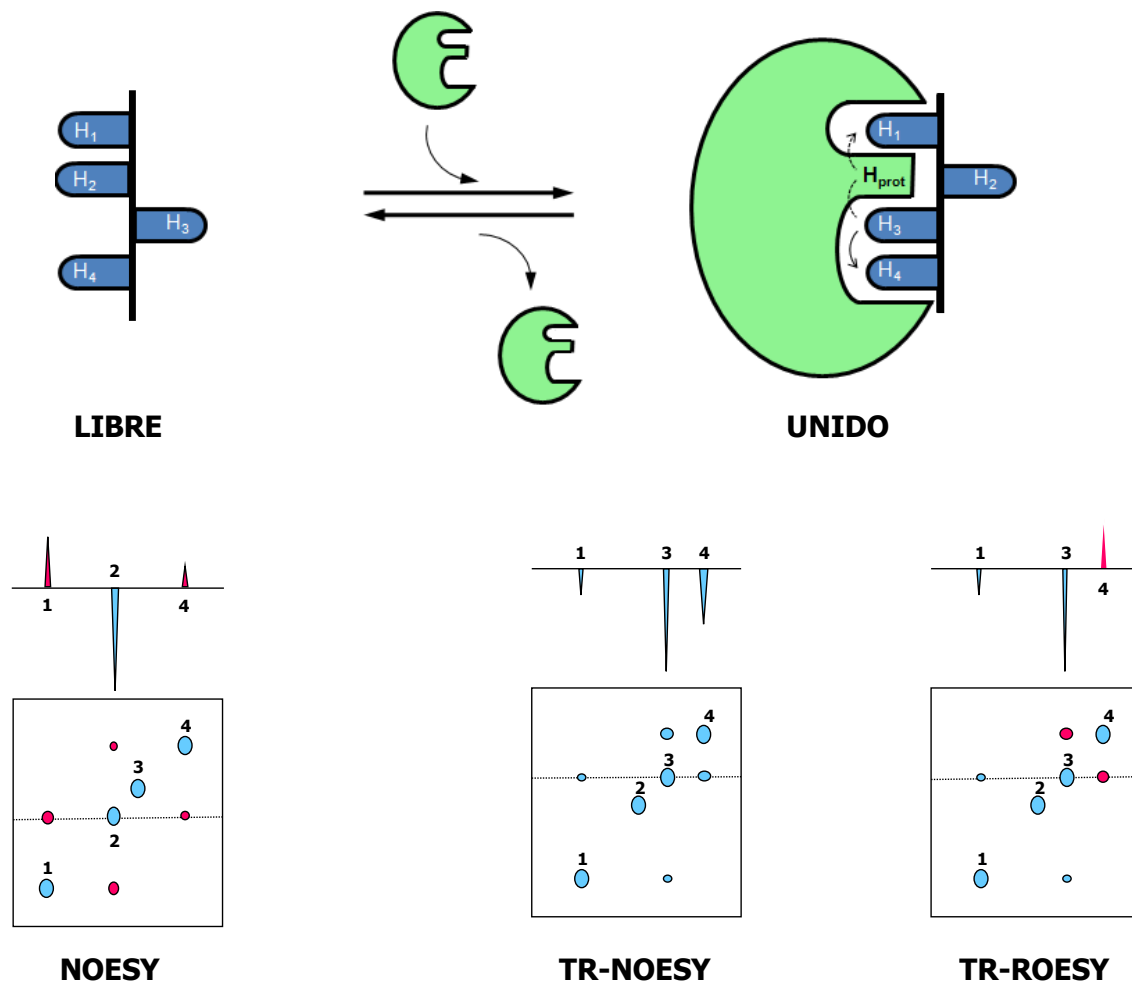


Figura 24: Esquema de la información proporcionada por el TR-NOESY. En estado libre, los picos de cruce de esta molécula son de distinto signo que los picos de la diagonal (*picos rosados*). En estado unido se pueden hacer dos distintos espectros, TR-NOESY y TR-ROESY. Todos los picos de cruce del TR-NOESY son del mismo signo que la diagonal (*color cian*). En cambio, el TR-ROESY permite diferenciar entre los contactos directos que existen en la conformación unida a la macromolécula (picos negativos, *cian*), y los contactos indirectos, en los que la transferencia de magnetización es mediada por algún protón de la proteína (picos positivos, *rosado*).

1.5.1.4. Secuencias Carr-Purcell-Meiboom-Gill (CPMG)

Estos experimentos permiten medir el tiempo de relajación transversal de un espín, T_2 . En un espectro 1D la anchura de línea de las señales depende de este parámetro. Las

señales estrechas y agudas de las moléculas pequeñas, con tiempos de correlación rápidos, corresponden a valores de relajación lenta (s), mientras que las señales anchas y romas están asociadas a macromoléculas, con tiempos de correlación lentos y tienen una relajación rápida (ms). El experimento CPMG, cuya adquisición está precedida por la repetición de un esquema *spin-eco*, produce señales de menor intensidad para los espines cuya relajación es rápida que para aquellos que muestran relajación lenta. Es decir, los tiempos de correlación están relacionados con las intensidades de las señales en los experimentos CPMG. En un proceso de reconocimiento molecular, el ligando adopta las propiedades de la macromolécula con la que interacciona. Por tanto, la intensidad de sus señales disminuirá. En una mezcla de compuestos se pueden identificar los ligandos reales, simplemente comparando dos CPMGs en ausencia y presencia de la macromolécula. Sólo se verán afectadas las señales de los ligandos.

Esta metodología ha sido utilizada principalmente usando ^{19}F -RMN¹⁴¹⁻¹⁴³. Su utilización en ^1H -RMN está limitada por la distorsión producida por la evolución de las J_{HH} durante los bloques de *spin-eco*. No obstante, este impedimento ha sido superado mediante la aplicación de distintas secuencias ad hoc.¹⁴⁴ De hecho, esta metodología se ha utilizado recientemente para identificar ligandos de oxigenasas (Figura 25)¹⁴⁵.

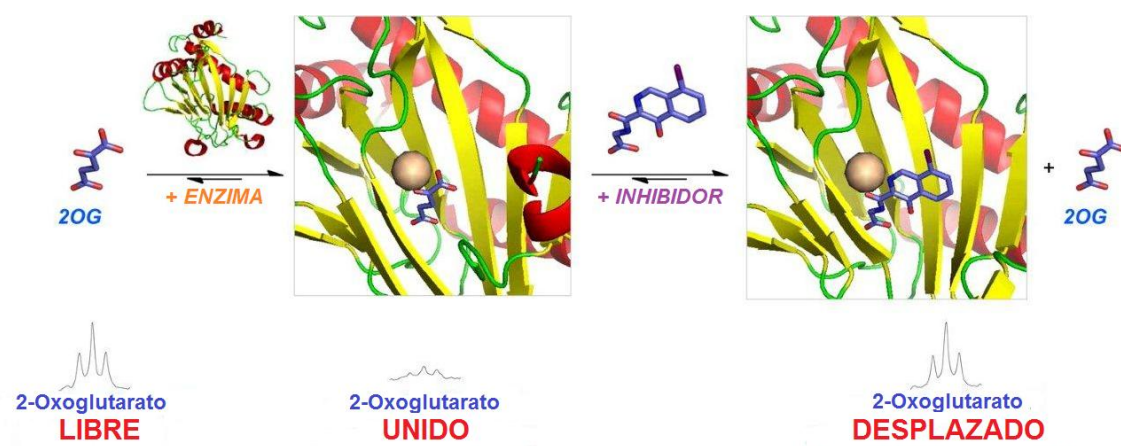


Figura 25: Detección de la interacción del 2-oxoglutarato frente a una oxigenasa mediante ^1H -CPMG. El reconocimiento molecular afecta a las señales del ligando (2OG). La presencia de un inhibidor que compite por el mismo sitio de unión determina que el ligando vuelva a comportarse como en su estado libre¹⁴⁵.

1.5.1.5. Experimentos de Relajación $T_{1\rho}$

La metodología y la interpretación de estos experimentos son parecidas a las utilizadas en las secuencias CPMG. En este caso, la secuencia de pulsos $T_{1\rho}$ utiliza un bloque de *spin-lock* antes de la adquisición, en vez del esquema *spin-eco* del experimento CPMG. Como en el caso anterior, el reconocimiento de los ligandos por parte de la macromolécula determina la disminución de la intensidad de sus señales.¹⁴⁶⁻¹⁴⁸

1.5.2. Métodos basados en cambios en la difusión translacional

El experimento DOSY (*Diffusion-Ordered SpectroscopY*)¹⁴⁹⁻¹⁵⁰ es un experimento bidimensional en el que una de las dimensiones muestra el desplazamiento químico y la otra, el coeficiente de difusión de las moléculas de la muestra. Es decir, correlaciona las señales de un compuesto con su coeficiente de difusión.

La secuencia DOSY contiene una repetición del esquema de *spin-eco* en presencia de gradientes. Los ligandos pequeños tienen una velocidad de difusión alta, característica de molécula pequeña, mientras que una proteína muestra una menor difusión translacional. La comparación de DOSYs en ausencia y en presencia de la proteína permite distinguir las moléculas que interaccionan de aquellas que no se unen¹⁵¹. De hecho, el análisis de experimentos DOSY con distintas relaciones molares ligando-proteína permiten calcular la constante de asociación del complejo correspondiente¹⁵².

Objetivos

En esta tesis doctoral se han estudiado cuatro sistemas diferentes utilizando distintas técnicas de RMN. Los dos primeros estudios analizan la estructura de péptidos, que están implicados en el dolor, en diferentes condiciones experimentales, y correlacionan los cambios estructurales con la modulación de su actividad. Los otros dos capítulos analizan dos procesos de reconocimiento carbohidrato-proteína utilizando distintos experimentos de RMN.

A continuación, se definirán los objetivos específicos de cada uno de los capítulos:

- Estudio estructural de análogos del péptido opiorfina

Caracterización estructural de distintos análogos de opiorfina mediante experimentos de RMN.

Relacionar las diferencias estructurales de estos análogos con las actividades obtenidas mediante los ensayos enzimáticos de fluorescencia.

- Estudio estructural de análogos glicosilados de nociceptina

Caracterización estructural de distintos análogos glicosilados de nociceptina mediante experimentos de RMN.

Relacionar las diferencias estructurales de estos glicopéptidos con las constantes de inhibición calculadas mediante un ensayo de desplazamiento del radioligando.

- El uso de etiquetas difluoroacetamida como sonda para detectar interacciones carbohidrato-proteína

Evaluación de la etiqueta difluoroacetamida como sonda en interacciones proteína-carbohidrato mediante distintos experimentos de ^1H -RMN y ^{19}F -NMR.

Estudiar el modo de unión de los análogos de N,N'-diacetil quitobiosa y N,N',N''-triacetil quitotriosa que contienen estas etiquetas difluoroacetamida.

- Inhibidores de β -galactosidasa a través de modificaciones en anillo aglicónico

Análisis conformacional de dos tiodisacáridos, que únicamente se diferencian en la estereoquímica de la posición C2 del anillo de pentopiranos, tanto en estado libre como en estado unido. Estas moléculas son inhibidores de la β -galactosidasa de *E.coli*.

Estudiar mediante cálculos de *docking* molecular y dinámica molecular el modo de unión de estos dos inhibidores.

Relacionar los resultados de los análisis conformacionales con las diferencias en las constantes de inhibición de estos inhibidores frente a la β -galactosidasa.

2. Estudio estructural de análogos del péptido opiorfina.

Las encefalinas endógenas, Met-enkephalin y Leu-enkephalin, se unen a dos tipos de receptores opioides, μ (MORs) y δ (DORs), que se encuentran principalmente en el Sistema Nervioso Central¹⁵³. Las encefalinas tienen mayor afinidad por los DORs¹⁵⁴, mientras la morfina es más afín por los MORs¹⁵⁵. La concentración sináptica de estos compuestos es la que modula la respuesta analgésica. Por tanto, el flujo de exocitosis de las encefalinas y la degradación de estos péptidos, que son hidrolizados principalmente por dos Zn-metalopeptidasas, endopeptidasa neural o neprilisina (*neural endopeptidase* o *neprilysin*, NEP) y aminopeptidasa-N (*aminopeptidase N*, APN) son mecanismos que controlan la señal analgésica¹⁵⁶.

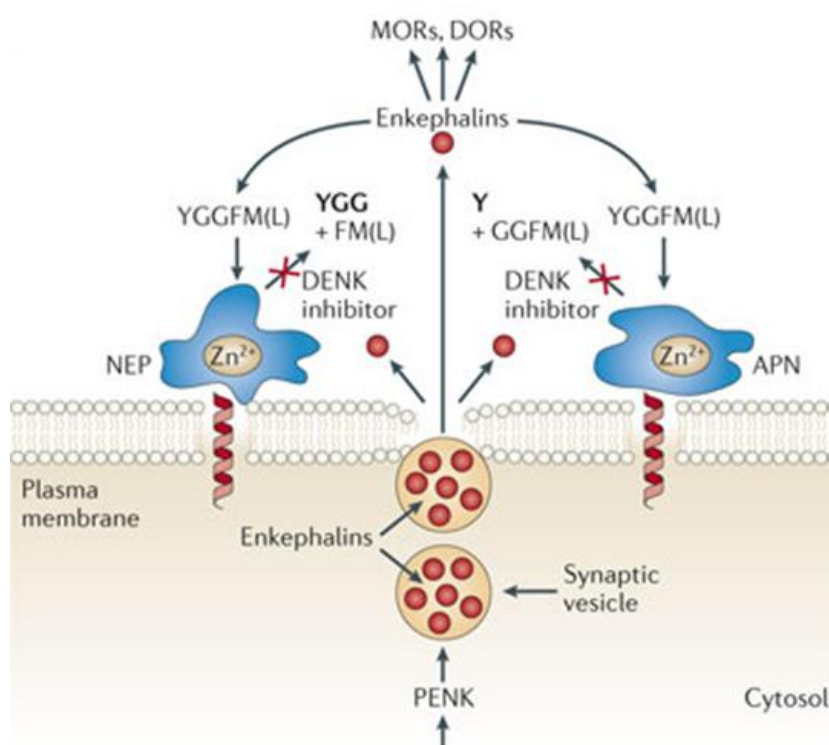


Figura 1: Las encefalinas (*enkephalins*) son péptidos sintetizados intracelularmente a partir del precursor preproencefalina (*preproenkephalin*, PENK). La exocitosis de las vesículas sinápticas que contienen estos péptidos es dependiente de Ca^{2+} . Cuando han sido liberadas, interaccionan con los receptores opioides (MORs y DOR), si su señal no es interrumpida por las Zn-metalopeptidasas, NEP y APN. La concentración de encefalinas modula la respuesta fisiológica analgésica.¹⁵⁶

Las secuencias de la *Met-enkephalin*, YGGFM, y de la *Leu-enkephalin*, YGGFL, se diferencian únicamente en el aminoácido C-terminal. La neprilisina, NEP, hidroliza las encefalinas en su enlace peptídico G-F, mientras que la aminopeptidasa, APN, hidroliza la secuencia Y-G. La degradación de estos neuropéptidos reduce el efecto analgésico, ya que disminuye la concentración sináptica de las encefalinas.

El esquema de actuación de las encefalinas¹⁵⁶ (Figura 1) señala dos estrategias de terapia, la utilización de moléculas exógenas que interaccionen con los receptores opioides o la inhibición de la hidrólisis de los ligandos endógenos. En este capítulo se va a abordar la generación de analgesia a través de la segunda estrategia, la inhibición de las enzimas hidrolíticas mediante moléculas análogas de la opiorfina. La opiorfina (*opiorphin*) es un pentapéptido, de secuencia QRFSR, aislado de la saliva humana¹⁵⁷. El efecto analgésico de la opiorfina en modelos animales demuestra que el catabolismo de las encefalinas puede ser bloqueado con fines analgésicos¹⁵⁸. Este péptido actúa como inhibidor endógeno de NEP y de APN e induce efectos analgésicos del dolor equiparables a la potencia de la morfina en modelos animales¹⁵⁹⁻¹⁶⁰.

El diseño de inhibidores de estas dos Zn-metalopeptidasas se ha afrontado mediante dos estrategias diferentes, el diseño de inhibidores duales o el diseño de inhibidores selectivos frente a una de estas enzimas. En el diseño de inhibidores duales se han descubierto compuestos que se unen covalentemente a ambas peptidasas¹⁶¹⁻¹⁶⁴, pero también se han diseñado moléculas cuya inhibición es dual sin ser covalente¹⁶⁵⁻¹⁶⁷. Desgraciadamente ninguna de estas propuestas ha tenido una aplicación clínica en el tratamiento del dolor. La otra estrategia, la inhibición diferencial, es más complicada ya que ambas proteínas tienen una alta similitud estructural que dificulta el diseño de inhibidores selectivos¹⁶⁸⁻¹⁶⁹. Sin embargo, el tiorfan (Figura 2, *thiorphan* en inglés) ha sido descrito como un compuesto selectivo frente a NEP, y se comercializa como agente antidiarreico por su acción periférica como inhibidor de esta hidrolasa¹⁷⁰⁻¹⁷¹.

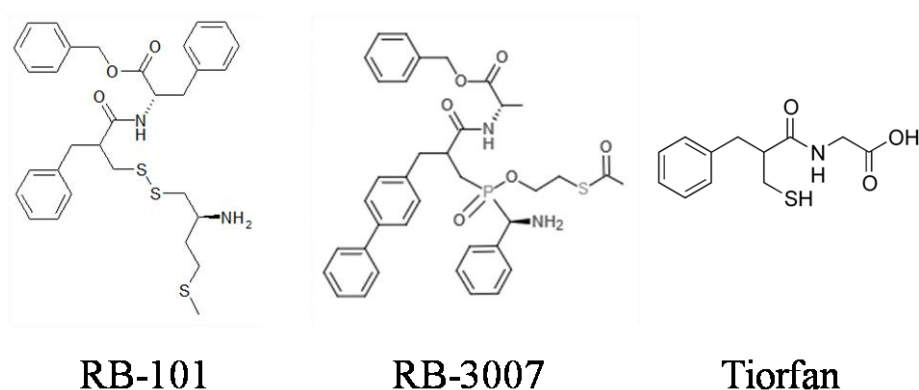


Figura 2: Estructura de algunos de los inhibidores de las Zn-metalopeptidasas, NEP y APN. RB-101¹⁶¹, inhibidor dual por unión covalente. RB-3007¹⁷², inhibidor dual de NEP y APN. Tiorfan¹⁷⁰, compuesto selectivo frente a NEP.

En este capítulo se ha analizado mediante RMN la opiorfina y una serie de análogos de este péptido natural. Los resultados del *barrido de alanina* han revelado que la fenilalanina tiene un papel crucial en la inhibición frente a NEP y a APN, así que se intentó modular la afinidad de este péptido introduciendo pequeñas modificaciones en este residuo aromático. A pesar de que la mayoría de los cambios no han mejorado la

actividad frente a ninguna de las enzimas, la sustitución del residuo natural L-Phe por su enantiómero D-Phe ha transformado este pentapéptido en un inhibidor selectivo de mayor potencia frente a APN. El análisis conformacional por RMN ha demostrado que estos cambios en actividad se correlacionan con un cambio conformacional en disolución.

2.1. Barrido de alanina, *Alanine scanning*, en la opiorfina

La importancia del residuo de fenilalanina en la actividad de la opiorfina fue descubierta mediante una sustitución sistemática de los residuos de la opiorfina por una Alanina (*Alanine scanning*)¹⁷³⁻¹⁷⁴. La sustitución de cada uno de los residuos de la opiorfina por alanina generó péptidos de menor afinidad para ambas dianas en prácticamente todos los casos. Cabe destacar que el cambio más drástico observado correspondió a la sustitución de la posición 3, la fenilalanina. Este análogo no mostró afinidad por ninguna de las dos metalopeptidasas (Tabla 1).

Péptidos	Secuencia	APN IC50 (μM)	NEP IC50 (μM)
Oporfina	NH ₃ ⁺ -Gln-Arg-Phe-Ser-Arg-COO ⁻	8.1	30
[Ala ¹]-Oporfina	NH ₃ ⁺ -Ala-Arg-Phe-Ser-Arg-COO ⁻	37	>100
[Ala ²]-Oporfina	NH ₃ ⁺ -Gln-Ala-Phe-Ser-Arg-COO ⁻	65	12
[Ala ³]-Oporfina	NH ₃ ⁺ -Gln-Arg-Ala-Ser-Arg-COO ⁻	>100	>100
[Ala ⁴]-Oporfina	NH ₃ ⁺ -Gln-Arg-Phe-Ala-Arg-COO ⁻	47	39
[Ala ⁵]-Oporfina	NH ₃ ⁺ -Gln-Arg-Phe-Ser-Ala-COO ⁻	>100	33

Tabla 1: Actividades frente a neprilisina (*NEP IC50*) y a aminopeptidasa-N (*APN IC50*) de la opiorfina y sus cinco análogos cuyos residuos han sido sistemáticamente sustituidos por alanina (*Alanine scanning*). Metodología descrita en la sección *materiales y métodos*.

2.2. Modulación de la actividad por sustitución del residuo de fenilalanina

Demostrada la importancia del residuo de fenilalanina en la actividad de la opiorfina, se procedió a un estudio singularizado de esta posición. Se diseñaron una serie de péptidos introduciendo diversas modificaciones, que pudieran modular la afinidad o la selectividad por alguna o ambas metaloproteasas, pero manteniendo siempre un anillo aromático en este residuo. La *L-Phe* natural fue sustituida por distintos análogos fluorados en el anillo aromático, por su enantiómero *D-Phe* y por el residuo elongado en un CH₂, homo-Phe.

La actividad de la opiorfina y de los siete análogos obtenidos por variaciones del residuo de fenilalanina se determinó usando los mismos ensayos que demostraron la importancia de la Phe³. La mayoría de las modificaciones por fluoración produjeron moléculas de menor actividad que la molécula inicial, en algún caso prácticamente inactivas. Por el contrario, el cambio de estereoquímica de este residuo produjo péptidos con una significativa selectividad. Se observó una mejora en la actividad sobre la aminopeptidasa-N (APN) en los compuestos en los que se sustituyó el aminoácido natural L-Phe por su enantiómero D-Phe. Tanto la D-Phe³-opiorfina como la [4-F]-D-Phe³-opiorfina mostraron mayor afinidad frente a la APN que la Oporfina. Además, esta actividad resultó selectiva para este receptor, ya que ninguno de los dos análogos demostró actividad en el ensayo fluorescente frente a neprilisina (Tabla 2).

Péptidos	Secuencia	APN IC50 (μM)	NEP IC50 (μM)
Oporfina	NH ₃ ⁺ -Gln-Arg-Phe-Ser-Arg-COO ⁻	8.1	30
[2-F]Phe ³ -Oporfina	NH ₃ ⁺ -Gln-Arg-[2-F]Phe-Ser-Arg-COO ⁻	44	48
[3-F]Phe ³ -Oporfina	NH ₃ ⁺ -Gln-Arg-[3-F]Phe-Ser-Arg-COO ⁻	>100	>100
[4-F]Phe ³ -Oporfina	NH ₃ ⁺ -Gln-Arg-[4-F]Phe-Ser-Arg-COO ⁻	13.5	>100
[F ₅]Phe ³ -Oporfina	NH ₃ ⁺ -Gln-Arg-[F ₅]Phe-Ser-Arg-COO ⁻	>100	>100
D-Phe ³ -Oporfina	NH ₃ ⁺ -Gln-Arg-D-Phe-Ser-Ala-COO ⁻	0.5	>100
[4-F]-D-Phe ³ -Oporfina	NH ₃ ⁺ -Gln-Arg-[4-F]-D-Phe-Ser-Arg-COO ⁻	0.5	>100
Homophe ³ -Oporfina	NH ₃ ⁺ -Gln-Arg-Homophe-Ser-Ala-COO ⁻	54	70

Tabla 2: Actividades frente a neprilisina (*NEP IC50*) y frente a aminopeptidasa-N (*APN IC50*) de la opiorfina y sus siete análogos cuya fenilalanina ha sido sustituida por distintos aminoácidos no-naturales.

2.3. Análisis conformacional por RMN

El cambio del aminoácido L-Phe por su enantiómero D-Phe debe producir un cambio en la estructura de estos péptidos, que mejora su actividad frente a APN e impide la interacción con NEP. Los posibles cambios en su estructura fueron analizados mediante RMN. El análisis de los espectros 1D de protón mostró diferencias notables entre los 8 péptidos analizados. Los distintos derivados fluorados de opiorfina mostraron pequeñas diferencias en sus desplazamientos químicos. En cambio, hay considerables cambios entre los péptidos con aminoácido natural L-Phe y aquellos con D-Phe o un análogo fluorado.

La desviación de los desplazamientos químicos de los protones α de los aminoácidos muestra la tendencia de una cadena peptídica a formar una estructura en α-hélice o en lámina β. Mediante el cálculo del índice CSI (*Chemical Shift Index*), se detectan las

diferencias entre los desplazamientos químicos de los H α medidos experimentalmente y los valores estándar para una estructura desordenada (*random coil*).¹⁷⁵ El análisis comparativo entre las distintas estructuras sólo reflejó pequeñas diferencias entre la opiorfina y los análogos analizados. Uno de los cambios más acusados fue la disminución de la tendencia a hebra β de la Arg² (1 \rightarrow 0) en los análogos que contienen D-Phe en lugar de L-Phe. Esta evidencia experimental esta apoyada también por la disminución del valor de la ³J_{NH-H α} de la Arg². Dicha constante en el caso de [4-F] Phe³-Oporfina (PFF) es 6.85 Hz y en el análogo [4-F]-D-Phe³-Oporfina (PFD) es 5.45 Hz. En cualquier caso, estas diferencias son leves porque estamos comparando estructuras con mucha flexibilidad (Tabla 3).

CSI*	Oporfina	[2-F]Phe³- Oporfina (FF2)	[3-F]Phe³- Oporfina (MPF)	[4-F]Phe³- Oporfina (PFF)	[F₃]Phe³- Oporfina (PF5)	D-Phe³- Oporfina (DPN)	[4-F]-D- Phe³- Oporfina (PFD)	Homophe³- Oporfina (HPE)
Gln¹ (4,118)	4,01 (-1)	4,03 (0)	4,02 (-1)	4,00 (-1)	4,03 (0)	4,03 (0)	4,04 (0)	4,09 (0)
Arg² (4,118)	4,32 (1)	4,32 (1)	4,34 (1)	4,34 (1)	4,34 (1)	4,19 (0)	4,21 (0)	4,40 (1)
Phe³ (4,398)	4,71 (1)	4,76 (*)	4,73 (*)	4,68 (*)	4,73 (*)	4,82 (*)	4,78 (*)	4,27 (*)
Ser⁴ (4,248)	4,44 (1)	4,45 (1)	4,46 (1)	4,43 (1)	4,49 (1)	4,45 (1)	4,44 (1)	4,42 (1)
Arg⁵ (4,118)	4,16 (0)	4,23 (1)	4,25 (1)	4,21 (0)	4,21 (0)	4,28 (1)	4,28 (1)	4,28 (1)

Tabla 3: Índice CSI (*Chemical Shift Index*) de los H α de la opiorfina y de sus 7 análogos cuya Phe en posición 3 ha sido substituida por otros residuos no-naturales. Según sea la diferencia del desplazamiento del protón α del aminoácido observado respecto del desplazamiento tabulado para el mismo protón en una estructura *random coil* (mayor de 0,1ppm, entre 0,1 y -0,1 o menor de -0,1, se asignan valores de CSI de 1, 0, -1 respectivamente).¹⁷⁵

Residuo	Protón (ppm)	Opiorfina	[2-F]Phe ³ - Opiorfina (FF2)	[3-F]Phe ³ - Opiorfina (MPF)	[4-F]Phe ³ - Opiorfina (PFF)	[F ₅]Phe ³ - Opiorfina (PF5)	D-Phe ³ - Opiorfina (DPN)	[4-F]-D-Phe ³ - Opiorfina (PFD)	Homophe ³ - Opiorfina (HPE)
Gln ¹	H α	4,01	4,03	4,02	4	4,03	4,03	4,04	4,09
Gln ¹	H β 2	2,08	2,11	2,11	2,11	2,14	2,11	2,11	2,17
Gln ¹	H β 3	---	2,04	2,04	2,03	2,03	---	---	---
Gln ¹	H ϵ 21	7,67	7,68	7,68	7,66	7,72	7,66	7,67	7,69
Gln ¹	H ϵ 22	7,02	7,04	7,04	7,02	7,07	6,99	7	6,99
Gln ¹	H γ	2,26	2,25	2,28	2,28	2,34	2,38	2,38	2,45
Arg ²	NH	8,88	8,89	8,91	8,88	8,95	8,91	8,94	9
Arg ²	H α	4,32	4,32	4,34	4,34	4,34	4,19	4,21	4,4
Arg ²	H β s	1,72	1,71	1,73	1,75	1,77	1,49	1,53	1,82
Arg ²	H δ s	3,16	3,15	3,15	3,17	3,2	2,96	3,01	3,21
Arg ²	H ϵ s	7,21	7,22	7,22	7,2	7,25	7,1	7,15	7,26
Arg ²	H γ 2	1,56	1,57	1,57	1,56	1,62	1,13	1,18	1,66
Arg ²	H γ 3	1,49	1,5	1,5	1,5	---	0,9	0,98	---
Phe ^{3*}	NH	8,75	8,74	8,79	8,78	8,93	8,96	8,97	8,86
Phe ^{3*}	H α	4,71	4,76	4,73	4,68	4,73	4,82	4,78	4,27
Phe ^{3*}	H β 2	3,19	3,08	3,19	3,15	3,25	3,37	3,32	2,17
Phe ^{3*}	H β 3	3,03	3,23	3,04	3,02	3,18	2,9	2,9	---
Phe ^{3*}	HH2	---	---	---	---	---	---	---	2,78
Phe ^{3*}	HH3	---	---	---	---	---	---	---	2,69
Ser ⁴	NH	8,41	8,44	8,45	8,38	8,54	8,4	8,43	8,45
Ser ⁴	H α	4,44	4,45	4,46	4,43	4,49	4,45	4,44	4,42
Ser ⁴	H β s	3,84	3,84	3,85	3,84	3,81	3,86	3,84	3,83
Arg ⁵	NH	8,13	8,33	8,35	8,25	8,33	8,43	8,45	8,35
Arg ⁵	H α	4,16	4,23	4,25	4,21	4,21	4,28	4,28	4,28
Arg ⁵	H β 2	1,9	1,9	1,89	1,88	1,86	1,91	1,92	1,89
Arg ⁵	H β 3	1,72	1,77	1,76	1,76	1,75	1,77	1,78	1,74
Arg ⁵	H δ s	3,22	3,21	3,24	3,22	3,22	3,22	3,21	3,13
Arg ⁵	H ϵ s	7,29	7,27	7,27	7,26	7,28	7,26	7,26	7,19
Arg ⁵	H γ s	1,62	1,65	1,67	1,64	1,63	1,66	1,65	1,59

Tabla 4: Los desplazamientos químicos de la opiorfina y de los siete péptidos análogos de la opiorfina en su posición 3. Las muestras fueron preparadas en una disolución acuosa 90:10 H₂O/D₂O a una concentración de 2 mM. Los espectros de RMN fueron adquiridos a una temperatura de 278K (90:10 H₂O/D₂O). El residuo en posición 3 depende del péptido analizado (*). Los valores se referenciaron al singlete de la referencia interna TSP (0 ppm).

La diferencia más drástica entre los péptidos que contienen L-Phe³ y aquellos que contienen D-Phe³ es el apantallamiento de los H γ de la Arg². Si comparamos los dos péptidos cuya Phe³ tiene un flúor en posición *para*, los protones H γ 2 y H γ 3 de la [4-F]-D-Phe³-Opiorfina se han desplazado 0,38 y 0,58 ppm, respectivamente, en relación a estos mismos protones en la [4-F]-Phe³-Opiorfina, PFF (Figura 3). Este efecto puede ser explicado por la proximidad del anillo aromático de la Phe (Figura 4). Por tanto, los péptidos que poseen D-Phe en vez de L-Phe deben adoptar una estructura que coloque la nube π del anillo aromático junto a la cadena lateral de la arginina en posición 2. Este dato experimental se apoyó en el cálculo computacional usando métodos de mecánica molecular y búsqueda conformacional (*Macromodel*). A pesar de la flexibilidad de estos péptidos, se encontró una restricción conformacional en los análogos con D-Phe, en los que el anillo aromático se situaba cerca de la cadena lateral de la Arg².

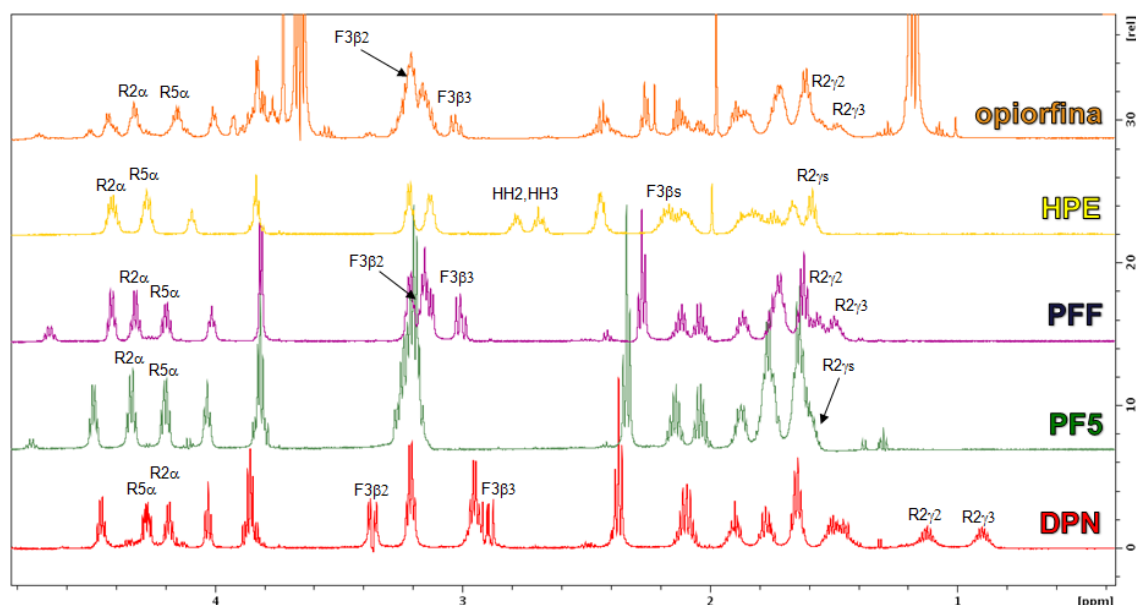


Figura 3: Superposición de los espectros 1D de protón de cinco de los péptidos ensayados. El espectro 1D de [4-F]-D-Phe³-Opiorfina, PFD, es similar al espectro de D-Phe³-Opiorfina, DPN, y los espectros de [2-F]-Phe³-Opiorfina, FF2, y de [3-F]-Phe³-Opiorfina, MPF, son similares al espectro de [4-F]-Phe³-Opiorfina, PFF. Los espectros están referenciados al TSP (0 ppm).

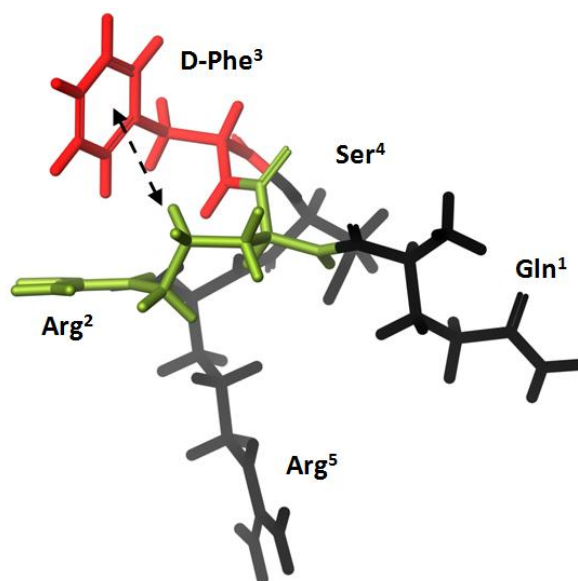


Figura 4: Estructura 3D del mínimo energético calculado (*materiales y métodos*) para la D-Phe³-Opiorfina, DPN. La proximidad entre los H γ de la Arg² (en color verde) y el anillo aromático de la D-Phe³ (en color rojo) ha sido resaltada con una flecha.

Asimismo, se comparó la orientación de la cadena lateral de la Phe para los análogos L-Phe y D-Phe. La rotación de la cadena lateral de la D-Phe mostró tres posibles mínimos, con valores de χ_1 de -60° , $+60^\circ$ y 180° . El ángulo χ_1 ($NH-C\alpha-C\beta-Carom$) puede ser caracterizado mediante las constantes de acoplamiento $^3J_{H\alpha-H\beta 2}$ y $^3J_{H\alpha-H\beta 3}$. El péptido DPN tiene unos valores de 4,6 Hz y 11,3 Hz, que están de acuerdo con los valores teóricos calculados para una conformación cuyo ángulo χ_1 es $+60^\circ$ o 180° (Figura 5). Pero la conformación $\chi_1=+60^\circ$ de la D-Phe es la única conformación que está de acuerdo con estos valores de constantes de acoplamiento y la proximidad del anillo aromático de la D-Phe a los H γ de la Arg² (Figura 4).

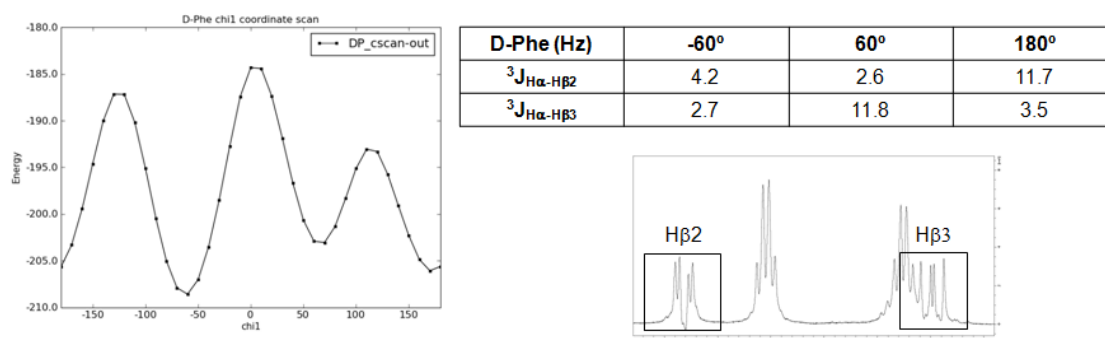


Figura 5: Análisis conformacional del aminoácido D-Phe mediante mecánica molecular. Se detectaron tres mínimos conformacionales para χ_1 , -60° , $+60^\circ$ y 180° . En la tabla se recogen los valores teóricos de las constantes de acoplamiento que poseen estos tres mínimos energéticos. Además, se muestra en detalle la región del espectro monodimensional de DPN donde aparecen ambos protones β de la D-Phe.

Se realizó el mismo análisis conformacional para el aminoácido natural, obteniéndose tres mínimos energéticos en torno a χ_1 . El análisis de las constantes de acoplamiento, $^3J_{H\alpha-H\beta}$, sólo pudo ser realizado para uno de los protones β debido al solapamiento de las señales. La $^3J_{H\alpha-H\beta3}$ observada es de 9,6 Hz, valor que no puede explicarse con la contribución de una única conformación (Figura 6). Por lo tanto, la opiorfina en medio acuoso no adopta una conformación única para el torsional χ_1 de la fenilalanina, como ocurre con su análogo DPN.

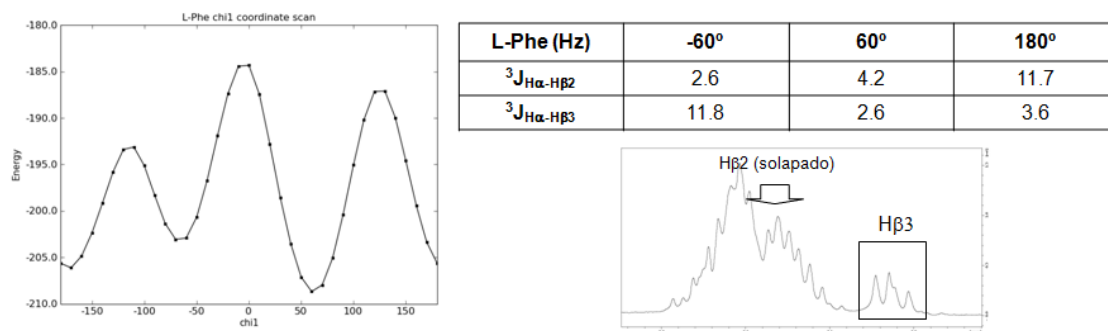


Figura 6: Análisis conformacional del aminoácido L-Phe mediante mecánica molecular. Se detectaron tres mínimos conformacionales con valores de χ_1 , -60°, +60° y 180°. En la tabla se recogen los valores esperados de las $^3J_{H\alpha-H\beta}$. Además, se muestra en detalle la región del espectro 1D de opiorfina donde aparecen los dos H β de la L-Phe.

Además de las restricciones observadas en torno a la cadena lateral de la D-Phe del análogo DPN, se identificaron contactos NOE que contribuyeron a la confirmación de la estructura propuesta por el análisis de mecánica molecular. El espectro NOESY de la D-Phe³-opiorfina mostró dos contactos interresiduales intensos, entre el NH de D-Phe³ y el NH de Ser⁴ (Figura 7, *Izda.*), y entre el NH de Arg⁵ y los H β s de Ser⁴ (Figura 9). Este patrón de NOEs del análogo de la opiorfina no se repite en el espectro NOESY de la opiorfina natural. En cambio, el péptido natural presenta un intenso NOE entre el NH de Arg⁵ y los H β s de Arg² (Figura 10), que prácticamente no se observa en el análogo que contiene la D-Phe³ (Figura 9). Además, se observan diferencias en las trazas de los protones γ de la Arg², los análogos que contienen los D-Phe sólo muestran contactos interresiduales entre estos protones apantallados y el grupo NH de la D-Phe (Figura 11), mientras que el análogo con L-Phe muestra contactos también con los NH de la Ser⁴ y la Arg⁵ (Figura 12). Estos distintos patrones de NOEs sugieren un cambio de la posición relativa de la Arg² y la Phe³, que modifica la superficie del potencial electrostático del péptido. Por lo tanto, la fuerte selectividad de los análogos que contienen la D-Phe por la aminopeptidasa-N puede ser explicada por este cambio en la posición relativa de sus aminoácidos (Figura 8). Publicaciones recientes han postulado que la flexibilidad del péptido natural, opiorfina, es el determinante de la promiscuidad de esta molécula por ambos mecanismos de hidrólisis de las encefalinas¹⁷⁶.

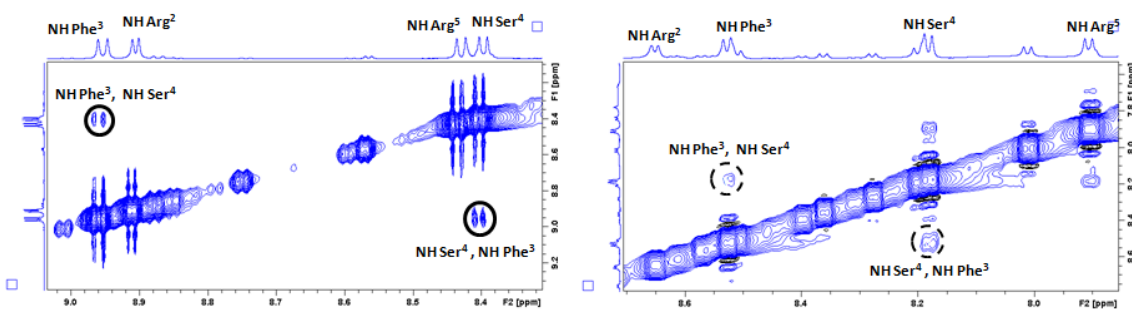


Figura 7: *Izda.*, detalle de la zona de los grupos NH en el espectro NOESY de la D-Phe³-Opiorfina, DPN. *Dcha.*, detalle de la zona de los grupos NH en el espectro NOESY de la opiorfina. Los espectros de RMN fueron adquiridos a una temperatura de 278K (90:10 H₂O/D₂O). Los espectros no están referenciados al TSP.

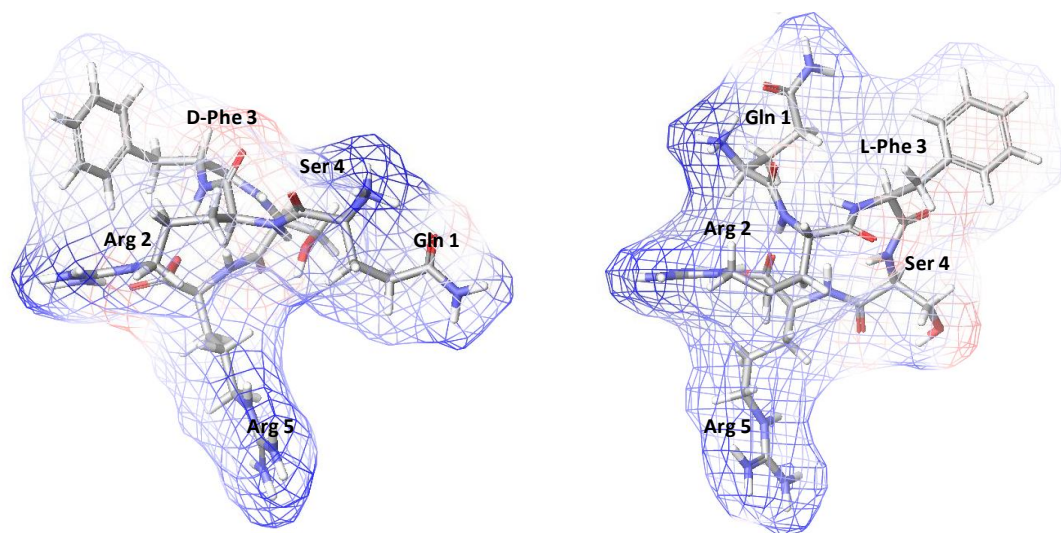


Figura 8: Representación de la superficie de potencial electrostático del análogo D-Phe³-opiorfina, DPN (*izda.*), y de la opiorfina (*dcha.*). La estructura representada de la opiorfina es el mínimo global de energía de la búsqueda conformacional. En el análogo D-Phe³-opiorfina la búsqueda conformacional se realizó con una restricción espacial entre el centroide de la D-Phe y los protones γ de la Arg². El conformero de mínimo energía de la búsqueda conformacional de este análogo es el que se ha representado en esta figura.

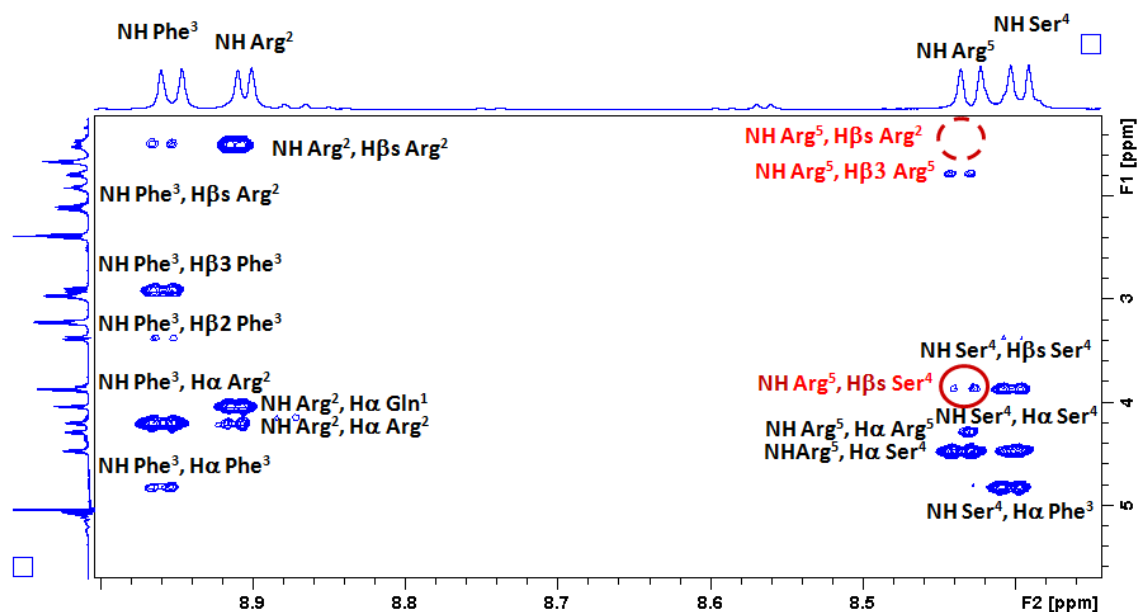


Figura 9: Detalle de la correlación dipolar de los grupos NH con el resto de protones de las cadenas laterales en el espectro NOESY de D-Phe³-Opiorfina, DPN. Los espectros de RMN fueron adquiridos a una temperatura de 278K (90:10 H₂O/D₂O). Los espectros no están referenciados al TSP.

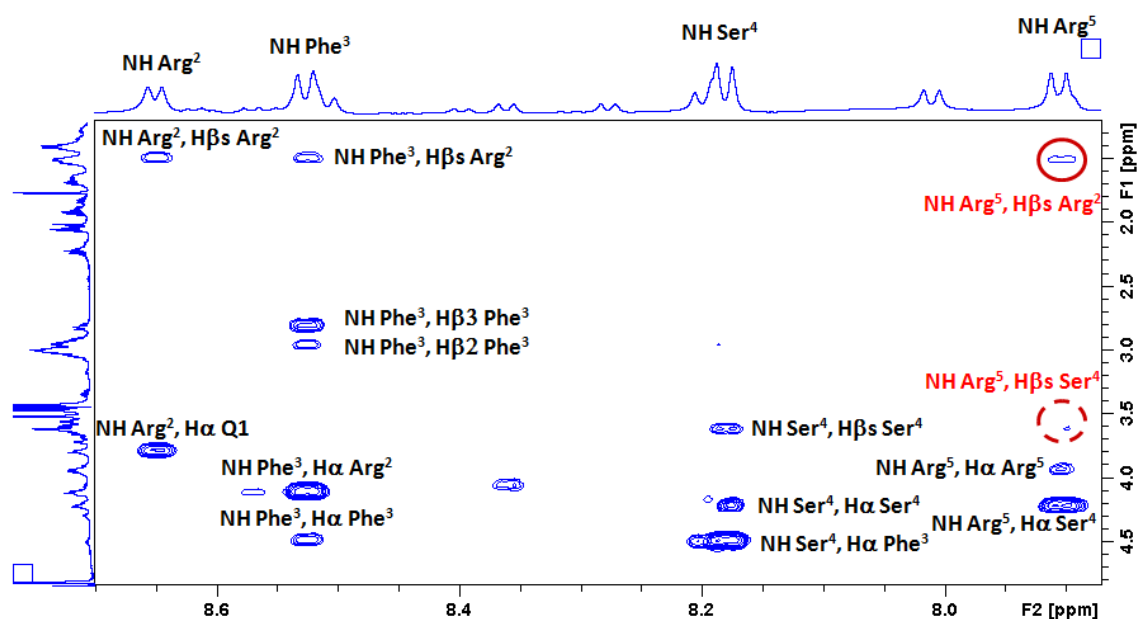


Figura 10: Detalle de la correlación dipolar de los grupos NH con el resto de protones de las cadenas laterales en el espectro NOESY de la opiorfina. Los espectros de RMN fueron adquiridos a una temperatura de 278K (90:10 H₂O/D₂O). Los espectros no están referenciados al TSP.

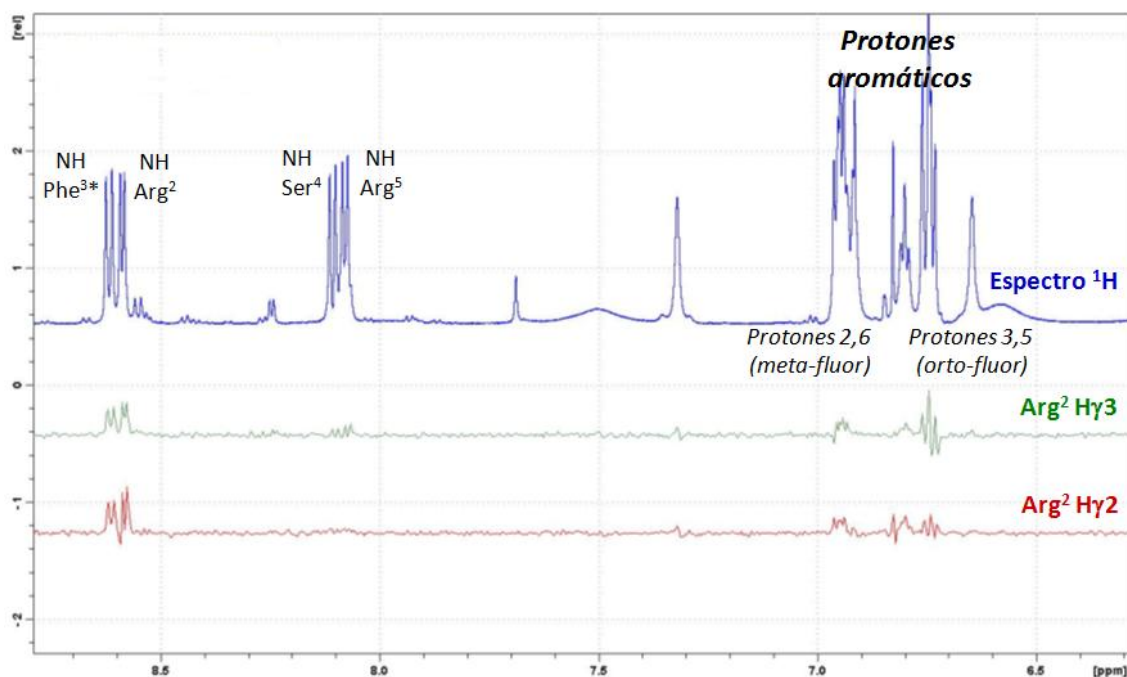


Figura 11: Superposición del espectro monodimensional de [4-F]-D-Phe³-Opiorfina, 4FD, y las trazas del espectro NOESY a partir de los protones γ de la Arg² sobre la diagonal. Los espectros no están referenciados al TSP.

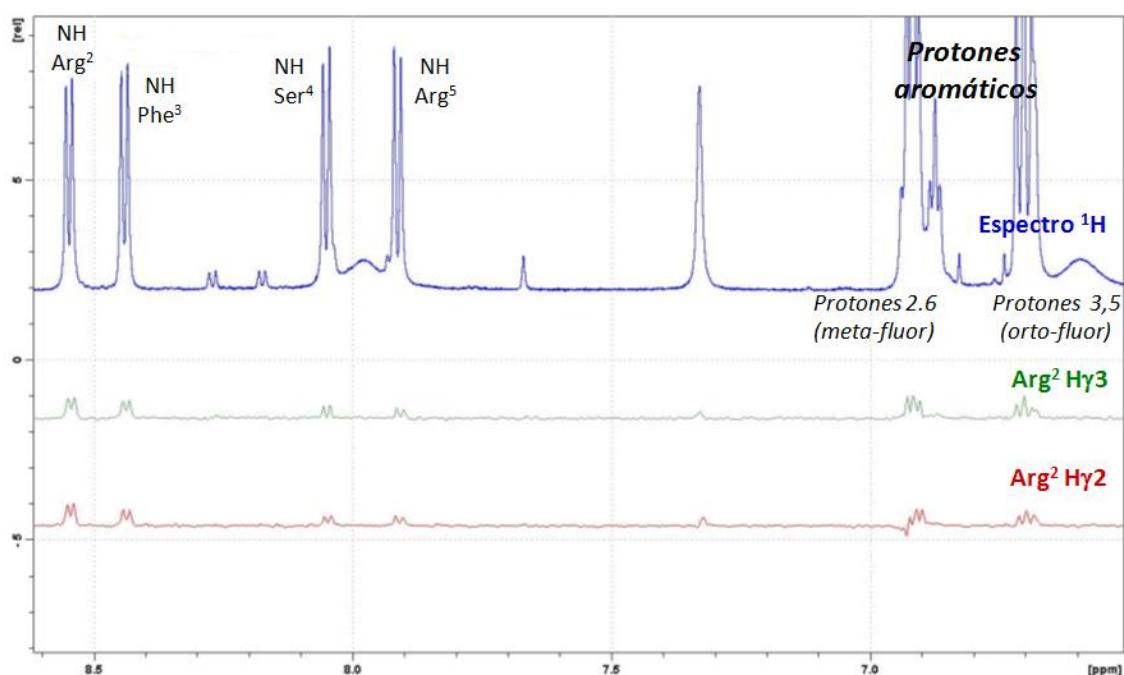


Figura 12: Superposición del espectro monodimensional de [4-F]-Phe³-Opiorfina, PFF, y las trazas del espectro NOESY a partir de los protones γ de la Arg² sobre la diagonal. Los espectros no están referenciados al TSP.

2.4. Conclusiones: relación estructura-actividad

En este capítulo se han estudiado distintos análogos de la opiorfina mediante RMN y se han encontrado evidencias espectroscópicas que correlacionan los cambios en actividad con una preferencia estructural de algunos de estos análogos. En primer lugar, el *barrido de alanina* en la opiorfina demostró la importancia del residuo de su posición 3, Phe³, en la unión de este péptido a las enzimas NEP y APN. Posteriormente, diversas modificaciones en el residuo aromático mostraron principalmente análogos que disminuían su actividad frente a ambas Zn-metalopeptidasas, pero los dos análogos que sustituían este aminoácido natural por su enantiómero, D-Phe, mejoraban su inhibición frente a APN y anulaban su actividad frente a NEP. Los estudios conformacionales por RMN mostraron diferencias estructurales entre los análogos que contienen D-Phe y aquellos que poseen L-Phe. La mejora en actividad frente a APN es debida a una selección conformacional en los péptidos que contienen D-Phe. Una posible interacción catión/ π entre el anillo aromático y el motivo guanidinio de la Arg² favorece la disposición espacial reconocida por la enzima APN. Este estudio ha identificado una estrategia para mejorar la potencia frente a APN mediante análogos que son metabólicamente más estables que el péptido natural.

2.5. Materiales y métodos experimentales

2.5.1. Síntesis de péptidos

La opiorfina y sus 7 análogos fueron producidos en el grupo Dr. Gregorio Valencia del IQAC-CSIC de Barcelona. La síntesis se realizó manualmente siguiendo una metodología estándar en fase sólida con prederivatizadas resinas Wang¹⁷⁷. Tras el acoplamiento manual de las distintas series de aminoácidos y la siguiente liberación de los péptidos de la resina, los péptidos fueron purificados con una columna de fase reversa en gradiente agua-acetonitrilo. Todos los péptidos fueron caracterizados mediante espectroscopía de masas de alta resolución, TLC, HPLC analítica y ¹H-NMR.

1.5.2. Cálculo de actividades frente a neprilisina (NEP) y a aminopeptidasa-N (AP-N)

El cálculo de las actividades frente a estas dos Zn-metalopeptidasas fue desarrollado por la Dra. Catherine Rougeot en el departamento de *Biochimie Structurale et Cellulaire* del Instituto Pasteur de París. Se usó un ensayo enzimático de fluorescencia para medir los valores de IC₅₀ de la opiorfina y de sus péptidos análogos frente a estas dianas. En el caso de la enzima NEP, se monitorizó la hidrólisis de un sustrato fluorescente específico para esta enzima: *Abz-dRGL-EDDnp FRET-peptide* (λ_{ex} a 320 nm; λ_{em} a 420 nm)¹⁷⁸. Para los cálculos de APN se utilizó *Ala-Mca* (λ_{ex} a 380 nm; λ_{em} a 460 nm)¹⁷⁹.

2.5.3. Resonancia Magnética Nuclear

Los experimentos fueron adquiridos en un equipo Bruker Avance 600MHz con una criosonda de triple canal (^{15}N -, ^{13}C -, ^1H -) a una temperatura de 278K. Los péptidos se disolvieron a una concentración de 2mM en una mezcla 90:10 $\text{H}_2\text{O}/\text{D}_2\text{O}$. Se utilizaron estas condiciones experimentales para ver con claridad las señales de los NH de las amidas del esqueleto peptídico y sus correlaciones con el resto de protones de la molécula. Al ser protones intercambiables se usaron temperaturas bajas para ralentizar el intercambio. La asignación se realizó usando experimentos 2D de correlación escalar (2D-TOCSY con tiempos de mezcla de 20 ms y 60 ms) y espacial (2D-NOESY con tiempos de mezcla de 300 ms). Los espectros se referenciaron al singlete (0 ppm) de la referencia interna TSP (2,2,3,3-tetradeterio-3-(TrimethylSilyl)Propionic acid).

2.5.4. Estructura tridimensional

La información estructural que se obtiene de los espectros de RMN se interpreta generando modelos 3D de las moléculas bajo estudio. En este trabajo se usó Macromodel 9.6¹⁸⁰ como programa de cálculo y de visualización. Las estructuras se contruyeron manualmente y seguidamente se realizó una minimización por mecánica molecular con el campo de fuerzas OPLS-2005¹⁸¹. Las coordenadas que se obtuvieron fueron el punto de partida de los cálculos de la búsqueda conformacional que se realizó con la función Maestro 8.5.110¹⁸² de Macromodel. El muestreo torsional de esta búsqueda estructural se realizó aplicando el método de Monte Carlo (*MCMM method*). Todos los cálculos se realizaron con el modelo de solvente implícito GB/SA (*continuum GB/SA solvent model*)¹⁸³ y con el método de minimización PRCG (*Polar-Ribiere conjugate gradient*).

3. Estudio estructural de análogos glicosilados de nociceptina

La nociceptina es un neuropéptido endógeno asociado con la percepción del dolor y con las perturbaciones del estado de ánimo¹⁸⁴⁻¹⁸⁶. Aunque está relacionado con los opioides, no muestra interacción con los receptores opioides clásicos de las subfamilias KOR, MOR y DOR (“Kappa-, Mu-, Delta- opioide receptor” respectivamente)¹⁸⁷. Este péptido es reconocido por otro tipo de receptor denominado NOP (*Nociceptin receptor*; o ORL-1, *opioid receptor like-1*). Como los receptores opioides clásicos, este receptor también pertenece a la superfamilia de receptores acoplados a proteínas G. Ambos, nociceptina y su receptor, se encuentran distribuidos principalmente en el sistema nervioso central¹⁸⁸. Como su propio nombre indica, este péptido y su receptor están implicados en la modulación de la nocicepción (percepción del dolor), y se ha encontrado que cuando el péptido se administra puede producir efectos tanto hiperalgésicos (mayor sensibilidad al dolor) como analgésicos dependiendo del lugar de aplicación¹⁸⁵.

La nociceptina, es un heptadecapéptido natural, cuya secuencia es FGGFTGARKSARKLANQ. Proviene de una preproteína (prepronociceptina de 176 aminoácidos)¹⁸⁹, y actúa como ligando endógeno de los receptores NOP. La señalización producida por este ligando produce una serie de efectos intracelulares descritos en otros receptores acoplados a proteínas G, como el bloqueo de canales de calcio, la modulación de quinasas o la activación de canales de potasio¹⁹⁰⁻¹⁹³. Esta pareja receptor-ligando muestra cierta selectividad¹⁹⁴. Otros ligandos de receptores opiáceos, como la naloxona, tienen poca afinidad por los receptores NOP; mientras que la nociceptina muestra baja afinidad por otros receptores opioides¹⁹⁵.

Aunque el reconocimiento y la señalización bioquímica se ha encontrado principalmente en el sistema nervioso central, sus efectos pueden reflejarse en una amplia serie de trastornos clínicos¹⁹¹. Estudios en el sistema nervioso central de mamíferos han demostrado su implicación en los sistemas del control del dolor. Curiosamente, la administración intracerebroventricular de este péptido determina un efecto antagónico a la analgesia de la morfina. En cambio, su administración por vía intratecal determina un efecto analgésico. Estas distintas respuestas muestran la complejidad de acción de esta pareja receptor-ligando en la regulación de los procesos del dolor¹⁹⁶.

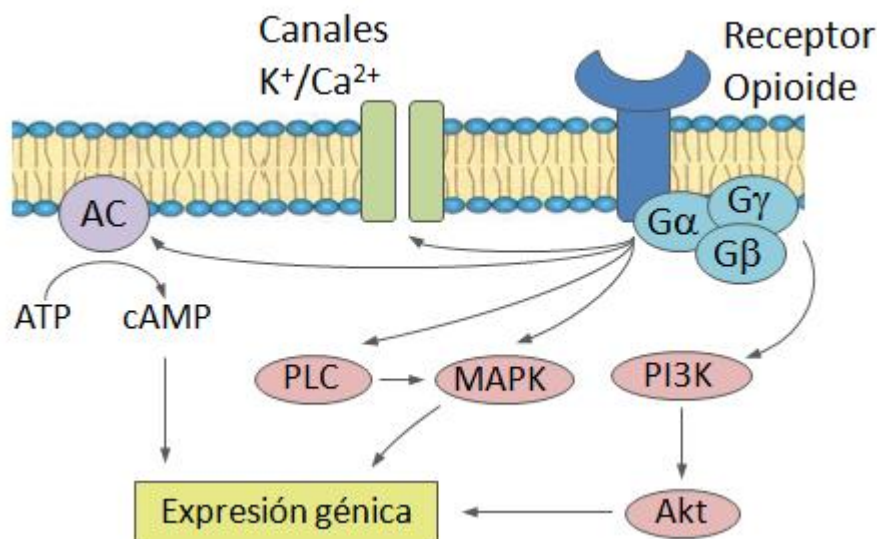


Figura 1: Esquema de las distintas rutas de señalización que provoca el reconocimiento de opioides como la nociceptina: estimulación de la adenilato ciclasa, la modulación de quinasas, bloqueo de canales de calcio o apertura de canales de potasio.

En esta sección se aborda el estudio estructural de la nociceptina y tres análogos glicosidados de este péptido mediante RMN. La glicosilación de un péptido produce cambios en su estructura que pueden afectar a la actividad de este péptido frente a su diana¹⁹⁷⁻¹⁹⁹. Este péptido se ha glicosilado en dos posiciones diferentes, en la Thr⁵ y en la Ser¹⁰. El péptido [Thr⁵-O- α -D-GalNAc]-nociceptina permitirá identificar la influencia de esta modificación en una zona más cercana del farmacóforo²⁰⁰, y las glicosilaciones en la Ser¹⁰, [Ser¹⁰-O- α -D-GalNAc]-nociceptina y [Ser¹⁰-O- β -D-GlcNAc]-nociceptina, ayudarán a explicar el efecto en una posición más alejada. Además, los dos análogos en la Ser¹⁰ permitirán analizar si hay distinción entre la configuración de la unión glicosídica, α o β . Los resultados estructurales y los datos de actividad de los tres glicopeptidos y el péptido natural ha permitido proponer un modo de acción para la nociceptina y mostrar la posibilidad de modular la actividad de estos péptidos del dolor mediante glicosilaciones en algunas posiciones de su secuencia.

3.1. Actividad de los péptidos frente a los receptores de membrana NOP

La actividad de los cuatro péptidos, la nociceptina y sus tres análogos glicosidados, fue calculada mediante un ensayo de desplazamiento del radioligando [³H]-nociceptina. Los péptidos no marcados radiactivamente, tanto el natural como sus tres análogos, se ensayaron a concentraciones crecientes sobre un homogenizado de membrana que contenía los receptores NOP y el radioligando. Se observó el desplazamiento del radioligando por los cuatro péptidos, y los datos se ajustaron con un modelo de

competición clásico 1:1. Según las constantes de inhibición, el compuesto más activo fue la [Ser¹⁰-O- α -D-GalNAc]-nociceptina (K_I 3.8 \pm 0.2 nM), seguido por la nociceptina natural (K_I 8.0 \pm 1.0 nM) y, con una constante similar a la nociceptina, la [Ser¹⁰-O- β -D-GlcNAc]-nociceptina (K_I de 10.4 \pm 2.6 nM). El compuesto que presentó menor actividad fue la [Thr⁵-O- α -D-GalNAc]-nociceptina (K_I 54.9 \pm 9.8 nM).

La glicosilación ha modificado la actividad de la nociceptina sobre los receptores de membrana NOP. No sólo afecta la posición de la glicosilación, sino que también lo hace el tipo de glicosilación y la naturaleza del glicano. El cambio más acusado se da entre los dos glicopéptidos que contienen α -GalNAc. De hecho, [Thr⁵-O- α -D-GalNAc]-nociceptina tiene una actividad un orden de magnitud inferior, mientras que [Ser¹⁰-O- α -D-GalNAc]-nociceptina mejora respecto a la nociceptina natural. La glicosilación del péptido en posición 10 con β -GlcNAc no modifica significativamente su afinidad frente a los receptores NOP.

Como primera hipótesis, la disminución de afinidad del glicopéptido [Thr⁵-O- α -D-GalNAc]-nociceptina podría estar relacionado con la perturbación del farmacóforo²⁰⁰. Se ha descrito que modificaciones sobre la posición 4 de la nociceptina, Phe⁴, presentan la capacidad de modular la actividad frente a los receptores de membrana²⁰¹. La glicosilación de la Thr⁵ posiblemente afecta la presentación y la orientación relativa de los aminoácidos. Además, en estudios previos del efecto de la glicosilación en construcciones truncadas de la nociceptina²⁰² se encontró que análogos truncados de nociceptina glicosilados en posición 5 mostraban menor actividad que aquellos truncados y glicosilados en la Ser¹⁰. Estos resultados están de acuerdo con los datos de actividad que han sido obtenidos para nuestros análogos, en los que se ha mantenido la cadena de la nociceptina completa. A continuación se realizó el estudio estructural de estos péptidos para determinar la correlación entre las diferencias en actividad observadas y la estructura de los respectivos péptidos glicosilados.

3.2. Dicroísmo circular de la nociceptina y sus análogos glicosilados

La caracterización estructural de los péptidos se inició estudiando su tendencia a formar α -hélice mediante el uso de dicroísmo circular (DC) de UV lejano. El espectro de DC de los cuatro péptidos se midió en agua, en una disolución de tampón fosfato a pH 7.4, y en trifluoroetanol (TFE). Este medio orgánico, de baja constante dieléctrica, facilita la estructuración de los péptidos, mimetizando el componente hidrofóbico de las membranas biológicas²⁰³⁻²⁰⁴.

Los perfiles de los espectros de los péptidos en agua y en tampón fosfato no mostraron estructura secundaria. Los espectros en estas dos condiciones experimentales solamente mostraron un efecto *Cotton* negativo a 200 nm (Figura 2 *arriba*). Este patrón de dicroísmo se correlaciona con cadenas peptídicas desordenadas. En cambio, los

espectros de estos péptidos en TFE mostraron tres efectos *Cotton* a distintas longitudes de onda: efecto *Cotton* positivo a 195 nm y efecto *Cotton* negativo a 205 nm y a 222 nm (Figura 2 *abajo*). Este perfil es característico de estructura secundaria de α -hélice²⁰⁵.

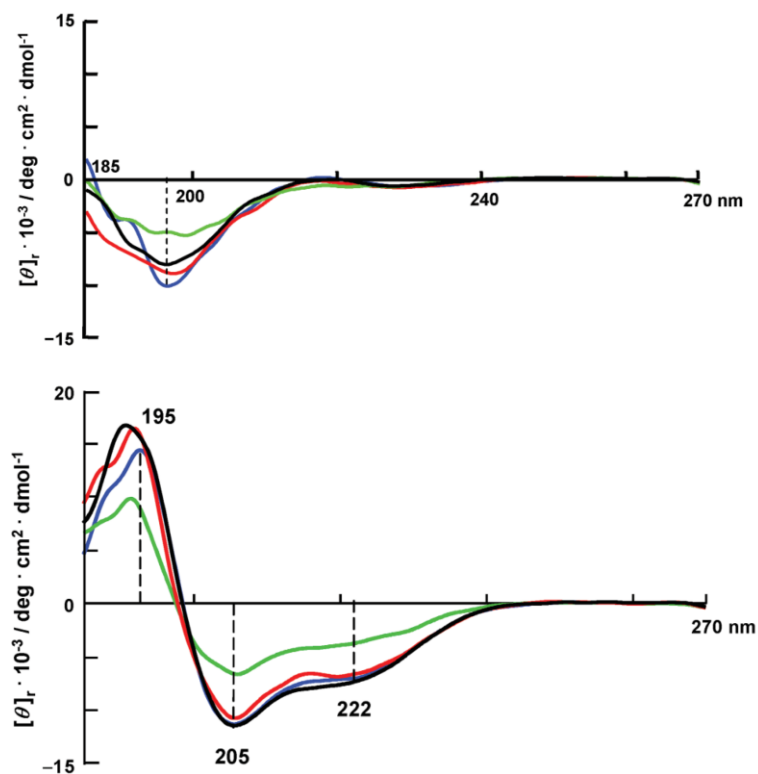


Figura 2: Espectros de DC del UV lejano de la nociceptina y sus tres análogos glicosilados. *Arriba*, superposición de los espectros de DC de los 4 péptidos en agua. *Abajo*, superposición de los espectros de DC en TFE. La nociceptina, en negro; [Ser¹⁰-O- α -D-GalNAc]-nociceptina, en azul; [Ser¹⁰-O- β -D-GlcNAc]-nociceptina, en verde; y [Thr⁵-O- α -D-GalNAc]-nociceptina, en rojo.

La *huella dactilar* de α -hélice se determina a partir de la relación entre el valor de elipticidad a 222 nm y a 205 nm. Se ha estimado que una cadena peptídica con una relación entre 0.8 y 0.95 tendrá un alto contenido en α -hélice²⁰⁵. La nociceptina y sus tres análogos glicosilados tienen una relación alrededor de 0.6, por lo que tienen una tendencia parcial a adoptar una estructura helicoidal en TFE. Si se quiere realizar una medida más cuantitativa de la contribución de α -hélice, se pueden aplicar distintas aproximaciones, que tienen en cuenta la elipticidad a 222 nm. En este estudio se ha utilizado el modelo de Luo y Baldwin para calcular la proporción de α -hélice de los péptidos. La nociceptina y sus dos análogos glicosilados con α -D-GalNAc, [Thr⁵-O- α -D-GalNAc]-nociceptina y [Ser¹⁰-O- α -D-GalNAc]-nociceptina, tienen un porcentaje de

α -hélice del 25%. En cambio, el análogo [Ser¹⁰-O- β -D-GlcNAc]-nociceptina tiene un porcentaje del 15%. Estos cambios en la helicidad de la cadena peptídica denotan la importancia del tipo de glicosilación, la introducción de una azúcar en configuración β en la posición 10, en vez de en α , disminuye la propensión de la cadena peptídica a formar una estructura helicoidal en presencia de TFE. Sin embargo, el decrecimiento en actividad de la [Thr⁵-O- α -D-GalNAc]-nociceptina no puede ser explicado únicamente por pérdida de estructura en TFE, ya que su porcentaje de α -hélice no es diferente del observado en la nociceptina natural y en la [Ser¹⁰-O- α -D-GalNAc]-nociceptina.

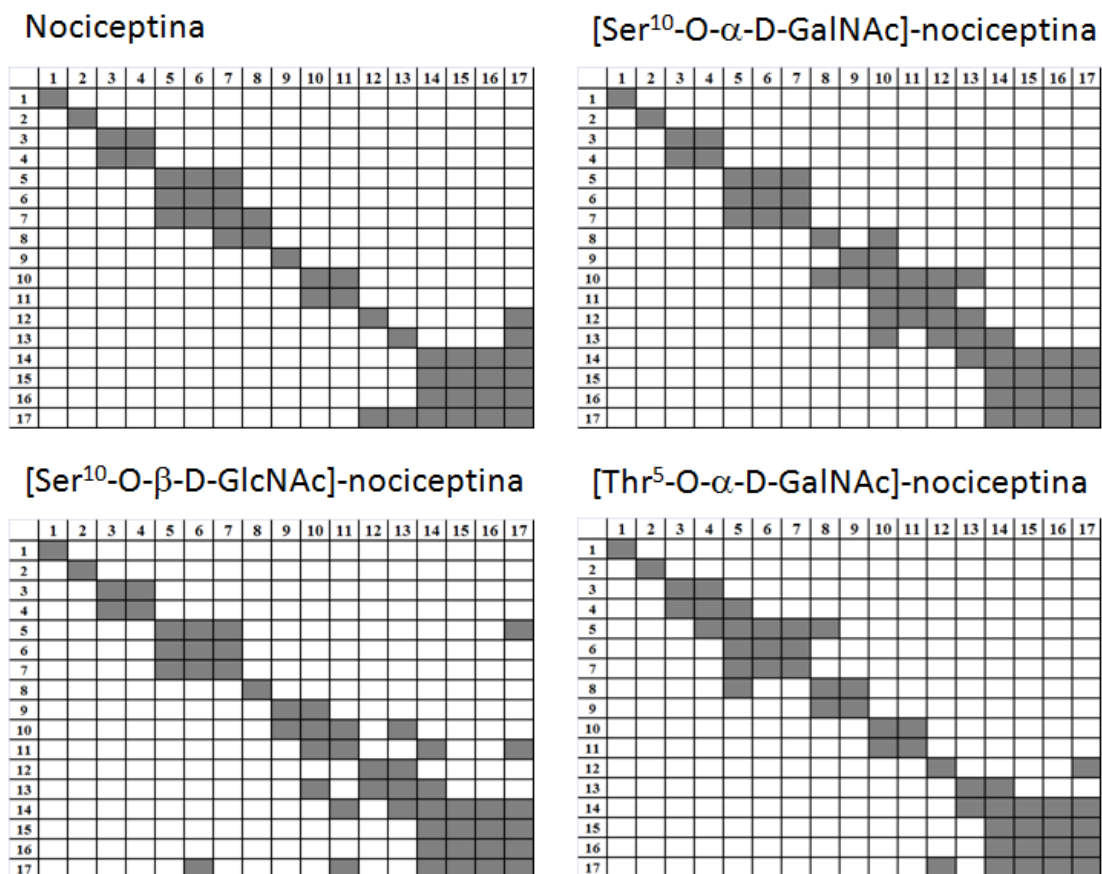
3.3. Resonancia Magnética Nuclear de la nociceptina y sus análogos glicosilados

El estudio estructural mediante RMN fue realizado en dos condiciones experimentales, en disolución acuosa y en presencia de micelas de SDS, con objeto de simular el ambiente de una membrana²⁰⁶⁻²⁰⁹.

En primer lugar se registraron una serie de espectros TOCSY y NOESY en ambas condiciones con objeto de asignar los desplazamientos químicos de los protones de cada uno de los péptidos siguiendo la metodología estándar mencionada en la introducción de esta tesis.

Seguidamente se analizaron los espectros NOESY. Los espectros en medio acuoso no mostraron ninguna evidencia de estructura ordenada para ninguna de las cadenas analizadas. Prácticamente no se detectaron picos NOE a larga distancia, sólo se observaron los contactos NOE secuenciales y algunos contactos interesiduales en la región carboxilo terminal. Los diagramas que se muestran en el esquema 1 resumen, de manera visual, los contactos NOE detectados para los cuatro péptidos. Además, las constantes de acoplamiento a tres enlaces, $^3J_{HH}$, entre los protones α y los protones amida de cada aminoácido mostraron unos valores intermedios en estas condiciones experimentales, indicando la existencia de promediados conformacionales y demostrando la disposición desordenada de estas cadenas.

La región amino terminal de los péptidos mostró mayor flexibilidad que el extremo carboxilo terminal²¹⁰. Los contactos observados en esta última región, sugieren cierta propensión helicoidal en disolución acuosa. La comparación cualitativa de los diagramas de los péptidos muestra un ligero incremento del número de NOEs en los análogos glicosilados. Sin embargo, los péptidos glicosilados no presentan cambios significativos en los desplazamientos químicos de sus señales, más allá de los esperados por la glicosilación en los residuos Thr⁵ y Ser¹⁰. La introducción del azúcar no afecta a la conformación desordenada de la nociceptina y sus análogos.

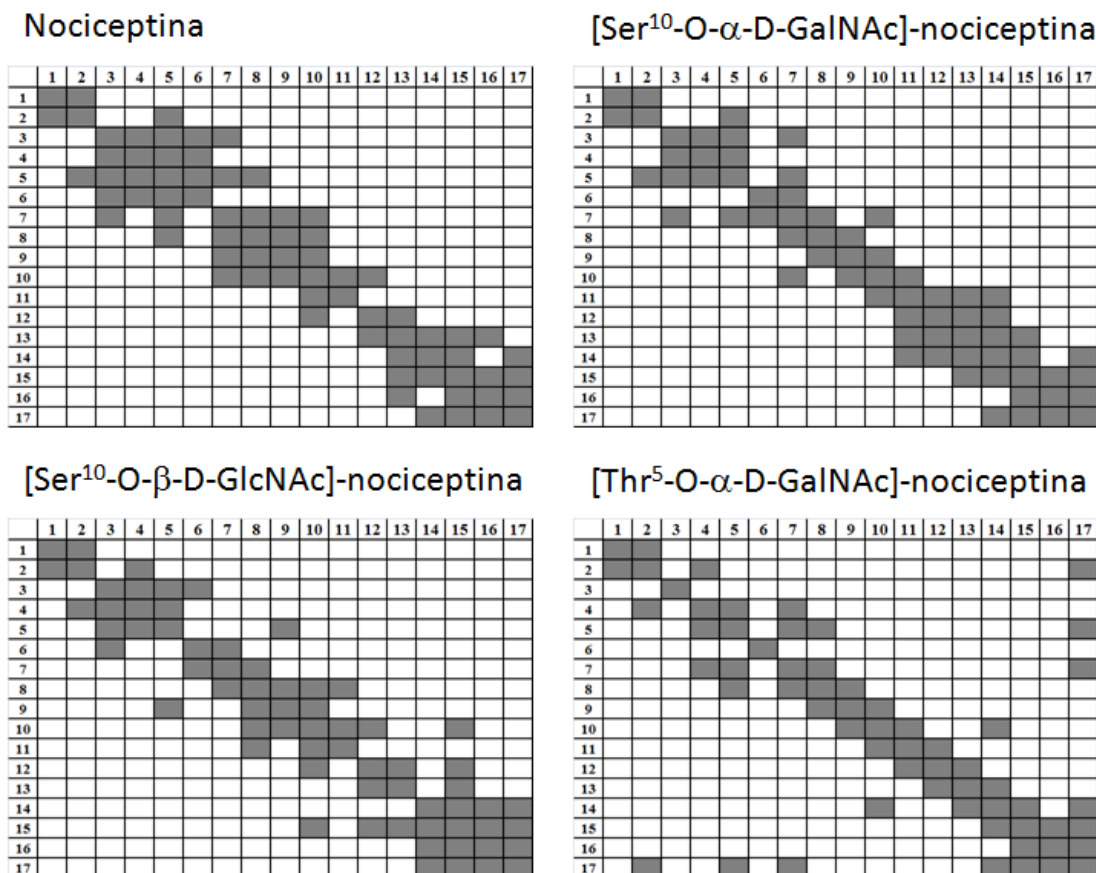


Esquema 1: Diagramas resumen de los contactos NOE observados para la nociceptina y sus análogos glicosilados en medio acuoso. Las celdas oscuras describen la presencia de contacto NOE interresiduales en la región de los protones NH entre los dos residuos que comparten celda. Los espectros NOESY fueron obtenidos a 278K y con 300 milisegundos de tiempo de mezcla.

A continuación se realizó el mismo procedimiento experimental en presencia de un medio mimético de membrana²¹¹⁻²¹³. Se adquirieron los espectros 2D NOESY en presencia de micelas de SDS. Una micela de SDS está compuesta por 60 moléculas de detergente²¹⁴⁻²¹⁵. Por lo tanto, las muestras de RMN se prepararon usando una relación molar SDS:péptido superior a 60:1. Además, utilizando estas condiciones experimentales se evitan las posibles interacciones intermoleculares entre dos moléculas de péptido en una misma micela.

Los contactos NOE interresiduales aumentan sustancialmente al compararlos con aquellos adquiridos en ausencia del medio micelar. La región de los protones amida muestra un mayor número de contactos NOE (Esquema 2). Los datos obtenidos para la nociceptina están de acuerdo con los publicados anteriormente²¹⁰. Estas primeras observaciones cualitativas denotan el aumento de la tendencia α -hélice de estas secuencias en presencia del medio mimético de membrana. Además, las estructuras helicoidales poseen valores pequeños de $^3J_{\text{NH-CH}\alpha}$, (Figura 4) lo que se refleja en la pérdida de las correlaciones entre sus protones amida y el resto de protones del sistema de espín en los experimentos TOCSY. En estas condiciones experimentales los

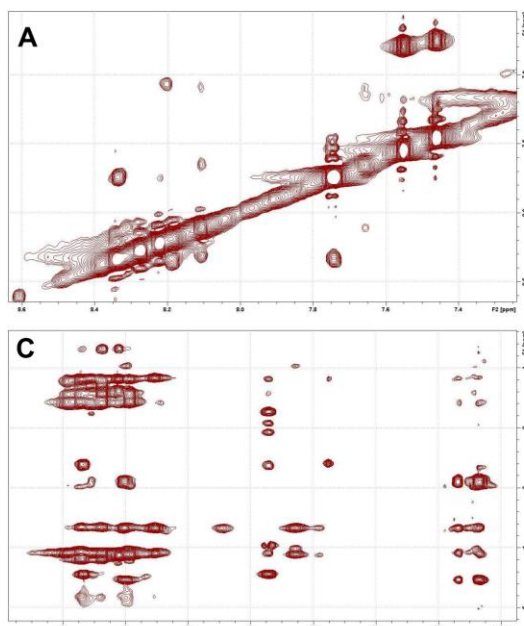
experimentos TOCSY tuvieron que ser adquiridos con mayores tiempos de mezcla (Figura 5).



Esquema 2: Diagramas resumen de los contactos NOE en la nociceptina y sus tres análogos glicosilados en presencia de micelas de SDS (en la región de los protones amida). Datos obtenidos mediante espectros NOESY de 300 ms a 288K.

La nociceptina tiene una triada básica de aminoácidos que se repite en la secuencia de este péptido, *Ala-Arg-Lys*. Esta triada está presente en las posiciones 7-9 y 11-13, así que hay que tener precaución al asignar los contactos NOE entre estos residuos. A pesar del solapamiento en estos espectros fue posible la asignación e identificación de los contactos NOE entre las dos triadas. Además, se identificaron inequívocamente los contactos entre el amino terminal de los péptidos y la triada 7-9, así como aquellos contactos que se dan entre el carboxilo terminal y la triada 11-13. Las dudas en la asignación de los picos de cruce NOE se excluyeron de la lista final de contactos para evitar errores que provinieran del solapamiento de las cadenas laterales de los aminoácidos de estas dos triadas básicas.

Nociceptina en medio acuoso



Nociceptina en presencia de SDS

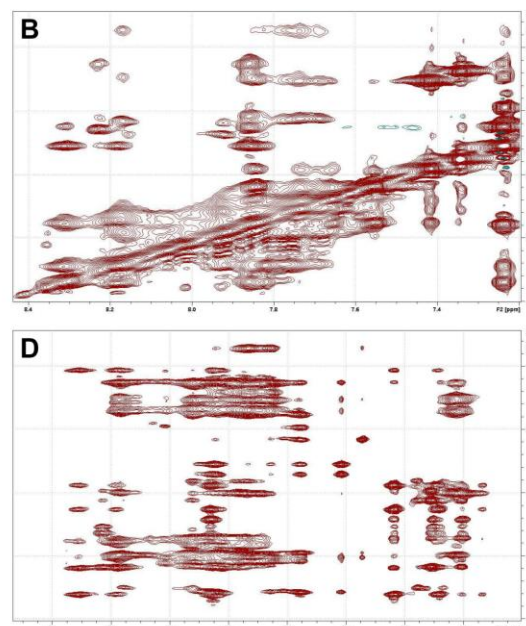


Figura 3: Detalle de los espectros NOESY de la nociceptina en medio acuoso y en presencia de micelas de SDS. Se muestran las regiones de correlación entre los protones amida (figuras A y B) y las regiones de correlación entre los protones amida y los protones alifáticos (figuras C y D).

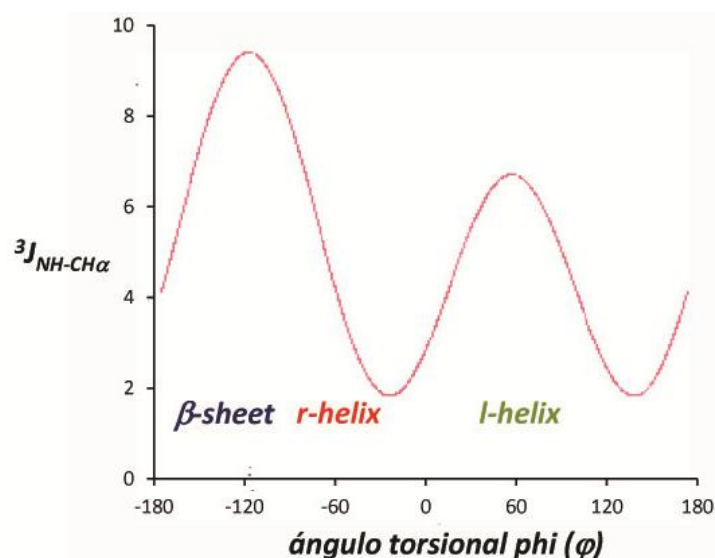
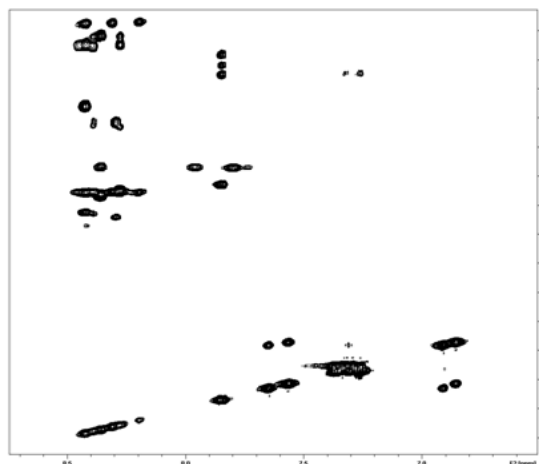


Figura 4: Relación entre el valor del ángulo ϕ y el valor de $^3J_{\text{NH-CH}\alpha}$. La estructura secundaria de una cadena peptídica se relaciona con los valores de $^3J_{\text{NH-CH}\alpha}$. β -sheet, lámina β (antiparalela $\phi=-139^\circ$, $^3J_{\text{NH-CH}\alpha}= 8,9\text{Hz}$; paralela $\phi=-119^\circ$, $^3J_{\text{NH-CH}\alpha}= 9,7\text{Hz}$). r -helix, hélice dextrógira (hélice $_{10} \phi=-60^\circ$, $^3J_{\text{NH-CH}\alpha}= 4,2\text{Hz}$; hélice $\alpha \phi=-57^\circ$, $^3J_{\text{NH-CH}\alpha}= 3,9\text{Hz}$). l -helix, hélice levógira ($\phi=57^\circ$, $^3J_{\text{NH-CH}\alpha}= 6,9\text{Hz}$).²¹⁶

Nociceptina en medio acuoso – TOCSY (60 ms)



Nociceptina en presencia de SDS – TOCSY (100ms)

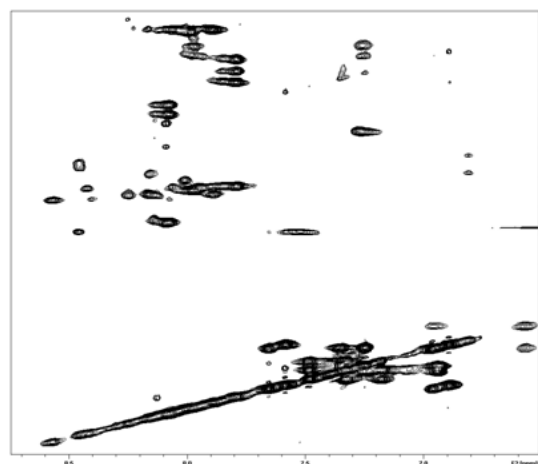


Figura 5: Detalle de los espectros TOCSY de la nociceptina en medio acuoso y en presencia de micelas de SDS. Los experimentos TOCSY de los péptidos en presencia de SDS tuvieron que ser adquiridos con mayores tiempos de mezcla (100ms) para poder observar las correlaciones en la región de los protones amida.

Los diagramas de tres de los péptidos, nociceptina, [Ser¹⁰-O- α -D-GalNAc]-nociceptina y [Ser¹⁰-O- β -D-GlcNAc]-nociceptina, muestran contactos que corroboran estructuras α -hélice en presencia de micelas de SDS. La figura 6 muestra la región del espectro NOESY de la nociceptina en la que se señalan los picos de cruce descriptivos de esta estructura. Estos picos de cruce se observaron también en los espectros NOESY de los dos análogos glicosilados en la Ser¹⁰. En esta región alifática se observan picos de cruce que muestran la hélice en el carboxilo terminal: contacto NOE entre uno de los H β de Asn¹⁶ y los H δ de Lys¹³, entre un H β de Asn¹⁶ y el H α de Lys¹³, entre el H α de Leu¹⁴ y los H β de Gln¹⁷, entre el H α de Leu¹⁴ y los H γ de Gln¹⁷, entre los H δ de Leu¹⁴ y los H γ de Gln¹⁷, entre los H γ de Arg¹² y el H α de Lys⁹, y, entre los H γ de Lys⁹ y los H δ de Arg¹².

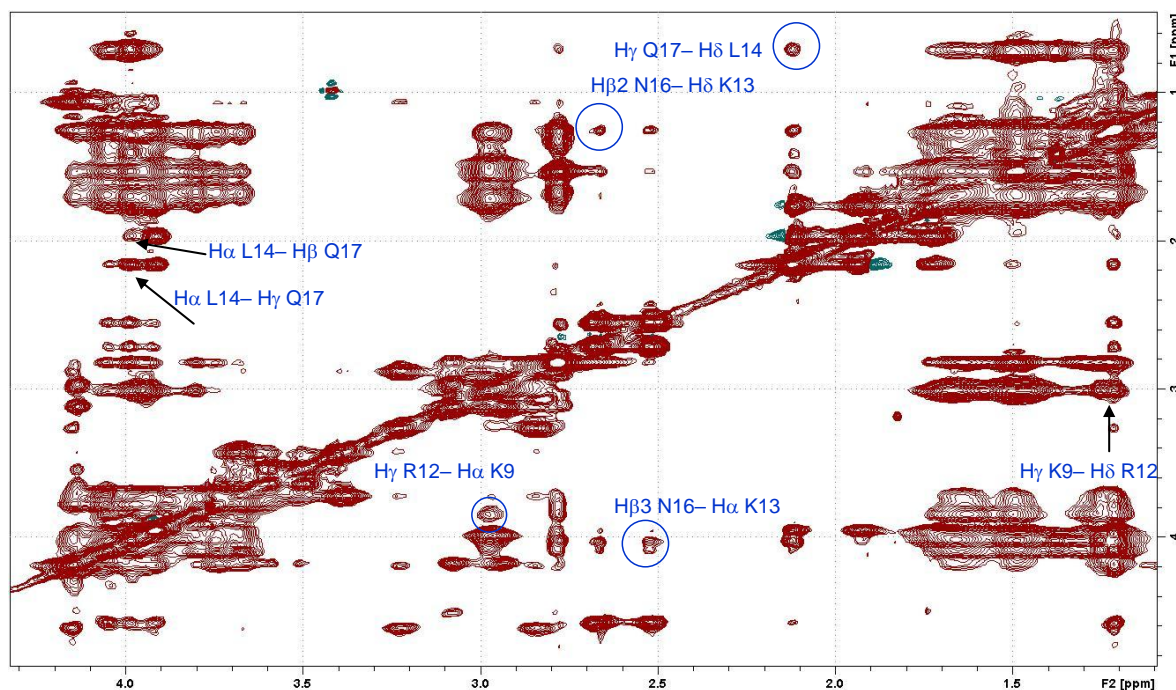


Figura 6: Región alifática de un espectro NOESY de Nociceptina (2 mM) en presencia de micelas de SDS (132 mM). Flechas y círculos azules señalan la localización de contactos interresiduales clave para justificar una estructura de α -hélice.

El péptido glicosilado en la Thr⁵ presenta la menor actividad de todos los estudiados aquí. En el diagrama de la [Thr⁵-O- α -D-GalNAc]-nociceptina (Esquema 2), este glicopéptido muestra un menor número de contactos NOE entre los protones amida de un residuo y los que pertenecen a residuos separados tres o cuatro aminoácidos en la secuencia. Esta diferencia refleja una menor tendencia de este glicopéptido a formar una estructura de α -hélice. Además, se observan ciertos contactos distales entre ambos extremos de la secuencia, entre los residuos 4, 5 y 7, y el aminoácido 17; en concreto, entre H ϵ Gln¹⁷ y todos los protones de la Thr⁵, entre H ϵ Gln¹⁷ y HN Ala⁷, y H ϵ Gln¹⁷ y H β Phe⁴ (Figura 7). Por lo tanto, este péptido tiene cierta tendencia a adquirir una estructura plegada. Los residuos de Thr O-glicosilados con α -GalNAc han mostrado tendencia a la formación de estructuras plegadas en modelos más sencillos²¹⁷⁻²¹⁸. Estos contactos distales no se presentan ni en nociceptina ni en los péptidos glicosilados en la Ser¹⁰ (Figura 7).

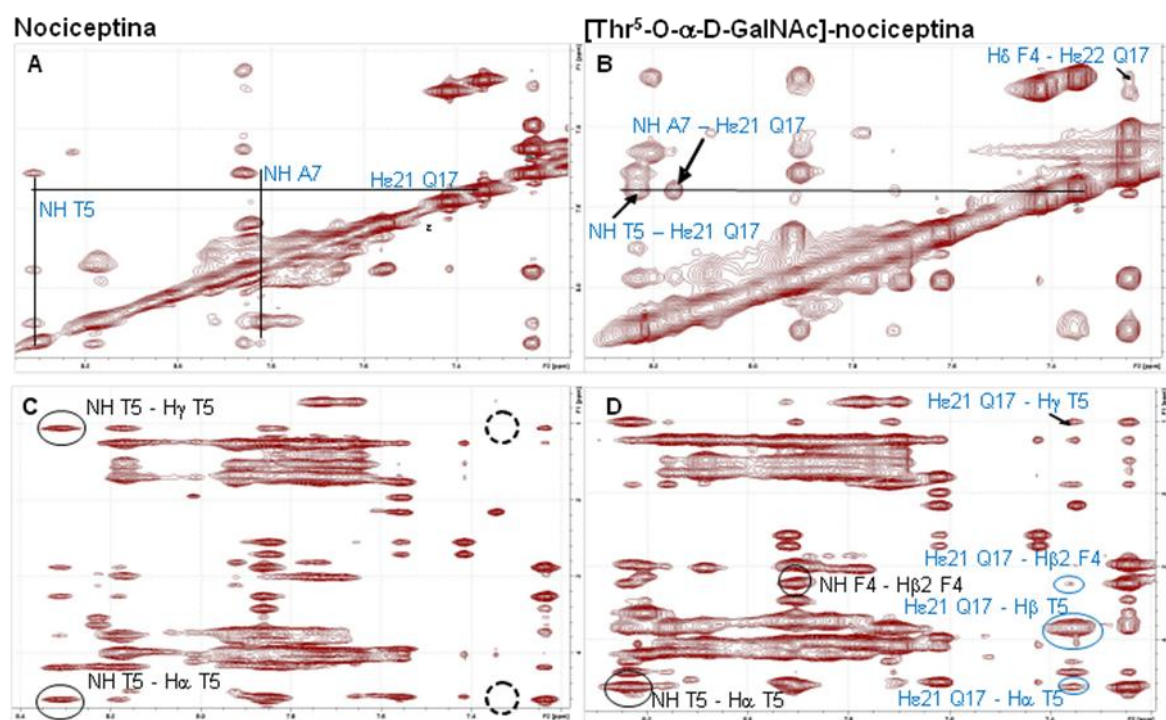


Figura 7: Regiones de los espectros NOESY en SDS para el péptido nociceptina (A y C) y su análogo glicosilado en Thr⁵ (B y D). Los picos de cruce intrarresiduales se indican en *negro*; los interresiduales, en *azul*.

Dada la existencia de un número considerable de picos de cruce interresiduales en los espectros NOESY en presencia de SDS, fue posible incluir esta información en protocolos de modelización molecular. Estos protocolos utilizan las restricciones experimentales para calcular la estructura tridimensional de péptidos y proteínas.

La obtención de modelos 3D se realizó usando el programa CYANA, que aplica un protocolo de cálculos de dinámica molecular torsional restringida por los NOEs experimentales. Se realizaron los cálculos introduciendo como restricciones experimentales 204, 188, 181, y 184 NOEs experimentales en los cálculos de la nociceptina, la [Ser¹⁰-O-α-D-GalNAc]-nociceptina, la [Ser¹⁰-O-β-D-GlcNAc]-nociceptina y la [Thr⁵-O-α-D-GalNAc]-nociceptina, respectivamente.

Además, se realizó un análisis de las desviaciones de los desplazamientos químicos de los protones alfa de cada aminoácido respecto de los correspondientes valores obtenidos en péptidos modelo lineales no estructurados (CSI: *chemical shift index*²¹⁹), como se ha descrito en el capítulo anterior. En el caso del péptido glicosilado en la ser 10 con α-GalNAc, las desviaciones observadas, fueron sistemáticamente negativas en la zona del carboxilo terminal indicativas de estructuras tipo hélice (Figura 8), confirmando nuevamente la tendencia de estos compuestos a adquirir una estructura helicoidal.

Existen trabajos previos que han postulado que la O-glicosilación y el tipo de sustitución, sobre Ser o Thr puede afectar a la estructura 3D de un péptido²²⁰. La

comparación de las estructuras 3D de la nociceptina y su análogo glicosilado [Ser¹⁰-O- α -D-GalNAc]-nociceptina no mostró grandes diferencias entre ambos compuestos. Ambos mantienen una estructura helicoidal, descrita ya previamente para el compuesto natural²¹⁰. Aunque existen pequeñas diferencias en el extremo carboxilo terminal, ya que el péptido glicosilado presenta una mayor tendencia de α -hélice. Diversos NOE interresiduales de los residuos que se encuentran entre Ala⁷ y Gln¹⁷, apoyados por los correspondientes CSI, corroboran esta estructura helicoidal. Además, el análisis de los ángulos ϕ y ψ mostró principalmente valores en la región asignada a α -hélice (gráficos de Ramachandran, Figura 9). Si prestamos atención a la región del aminoácido O-glicosilado (*Lys-Ser-Ala*), las 20 estructuras de menor energía conformacional propuestas por CYANA presentan valores de ángulos de torsión correspondientes a α -hélice.

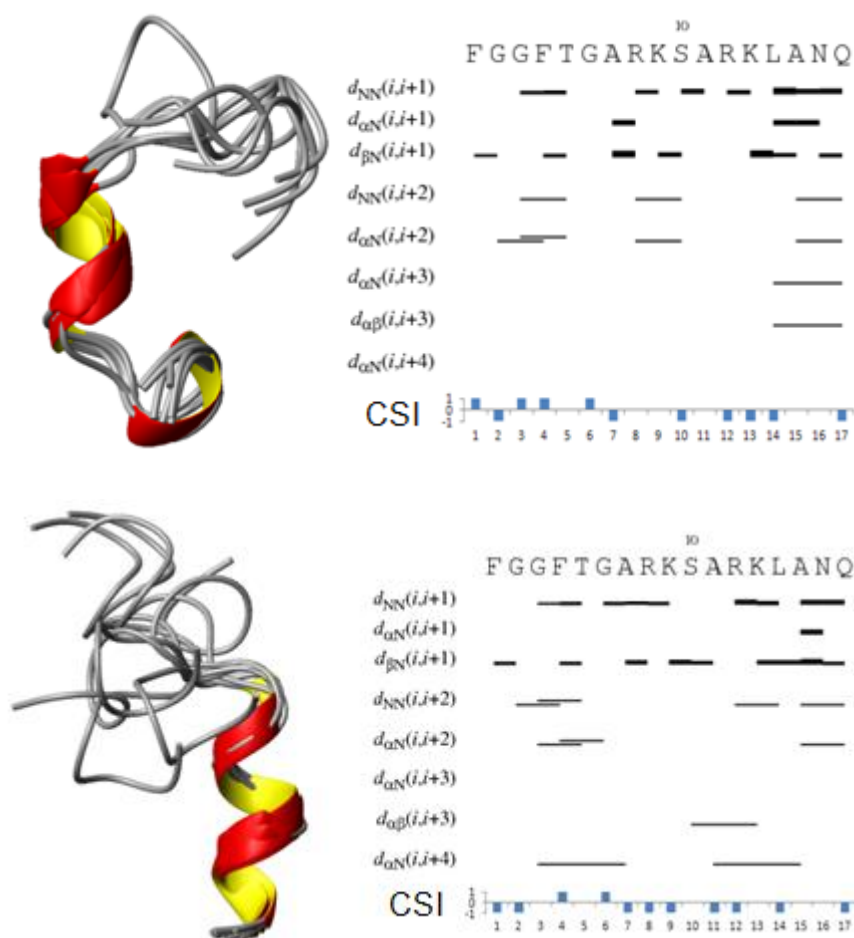


Figura 8: Comparación de las estructuras calculadas por CYANA para la nociceptina, *arriba*, y su análogo [Ser¹⁰-O- α -D-GalNAc]-nociceptina, *abajo*. Las estructuras son acompañadas por un resumen de los contactos NOE y sus CSI.

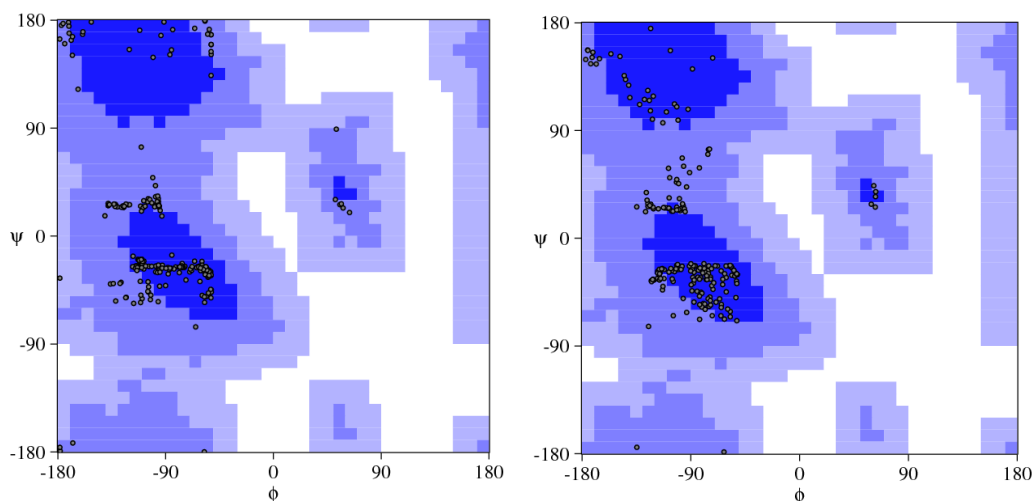


Figura 9: Mapas de Ramachandran de la nociceptina, *izquierda*, y su análogo [Ser¹⁰-O- α -D-GalNAc]-nociceptina, *derecha*. Los puntos corresponden a los valores de ϕ ψ para cada aminoácido de las 20 estructuras de CYANA.

A partir de las estructuras de CYANA, se determinó la disposición relativa de la α -D-GalNAc respecto a la cadena peptídica mediante el análisis de los contactos NOE entre el glucósido y los aminoácidos (Figura 10). Un análisis conformacional previo estableció los requerimientos de los ángulos de torsión χ_1 (N-C α -C β -O1), ϕ_g (O5-C1-O1-C β) y ψ_g (C1-O1-C β -C α) entre el residuo de serina y la N-acetil-galactosamina para cumplir los datos experimentales de NOE. El ángulo ϕ_g está restringido a 60° debido al efecto exo-anomérico y, ψ_g muestra cierta tendencia a presentar una disposición *anti*²¹⁷. En cambio, el ángulo χ_1 presenta mayor flexibilidad, con dos mínimos energéticos, $\chi_1 = 60^\circ$ y $\chi_1 = -150^\circ$ (Figura 10). De hecho, los picos de cruce del NOESY no pueden ser explicados exclusivamente usando un único mínimo conformacional, por lo que ambas disposiciones estructurales deben estar presentes en estas condiciones experimentales (Figura 11).

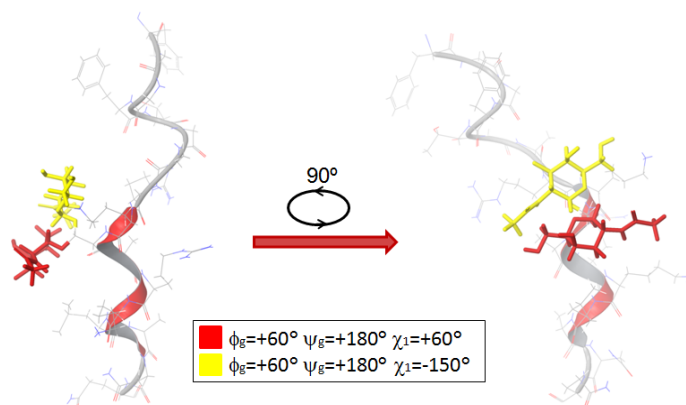


Figura 10: Modelos 3D del péptido [Ser¹⁰-O- α -D-GalNAc]-nociceptina. Se muestran superpuestas las dos disposiciones del α -D-GalNAc respecto a la cadena peptídica, $\chi_1 = 60^\circ$ (*rojo*) y $\chi_1 = -150^\circ$ (*amarillo*).

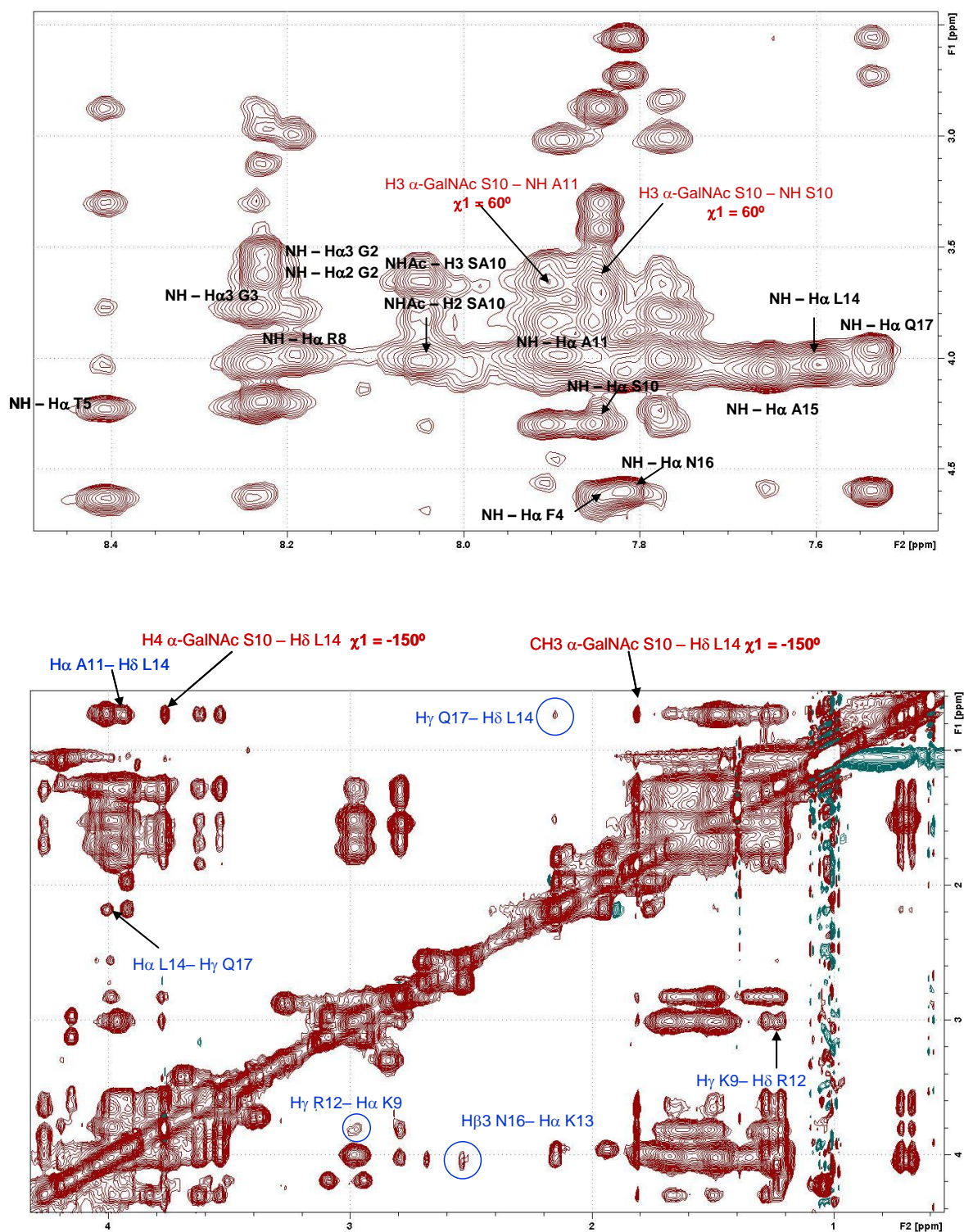


Figura 11: Detalle de los espectros NOESY del análogo glicosilado [Ser¹⁰-O- α -D-GalNAc]-nociceptina en presencia de micelas de SDS. Las etiquetas azules muestran los NOEs característicos de α -hélice. Las rojas marcan los NOE exclusivos de cada una de las disposiciones espaciales del residuo glicosilado, $\chi_1=60^\circ$ y $\chi_1=-150^\circ$. *Primer panel*, región de correlación HN/H α . *Segundo panel*, región de la correlación entre H_{alifáticos}.

Este mismo análisis fue realizado con los otros dos análogos. El glicopéptido [Ser¹⁰-O-β-D-GlcNAc]-nociceptina mostró mayor flexibilidad que el péptido natural en presencia de SDS. El NOESY muestra pocos picos de cruce característicos de una estructura helicoidal en su extremo carboxilo terminal, lo que determina que únicamente 3 de las 20 estructuras propuestas por CYANA tengan una estructura helicoidal claramente definida. Si prestamos atención a la triada de aminoácidos en torno a la glicosilación (Lys-Ser-Ala), los valores de ϕ y ψ de los enlaces peptídicos presentan una mayor dispersión que los que adoptan para los dos péptidos anteriores. Trabajos anteriores con sistemas más sencillos han demostrado que los tres ángulos de torsión en Ser-O-β-D-GlcNAc son más flexibles que la glicosilación correspondiente a Thr²¹⁸. Según estos estudios, los valores del ángulo glicosídico ϕ_g (O5-C1-O1-Cβ) en los que se cumple la condición del efecto exo-anomérico (60° y -60°) se combinan con valores poco variables del ángulo aglicónico ψ_g (C1-O1-Cβ-Cα). Además, se ha observado una alta flexibilidad en el ángulo χ_1 (N-Cα-Cβ-O1) para esta glicosilación. El análisis de la disposición relativa del residuo de GlcNAc frente a la cadena peptídica fue, en este caso, muy complicado, debido a la alta flexibilidad. Asumiendo unos valores de ϕ_g y ψ_g de 60° y 180°, respectivamente, se necesitan los tres valores de χ_1 para explicar los NOE experimentales. El NOE entre el grupo acetilo de la β-GlcNAc y el NH Ser¹⁰ está presente cuando χ_1 es -60°, el NOE entre el H4 β-GlcNAc, el NH de la Ser¹⁰ es intenso para el mínimo energético cuyo χ_1 es 60°, y el NOE entre el protón H4 de la β-GlcNAc y el metilo de la Ala7 se produce cuando su χ_1 es 180°.

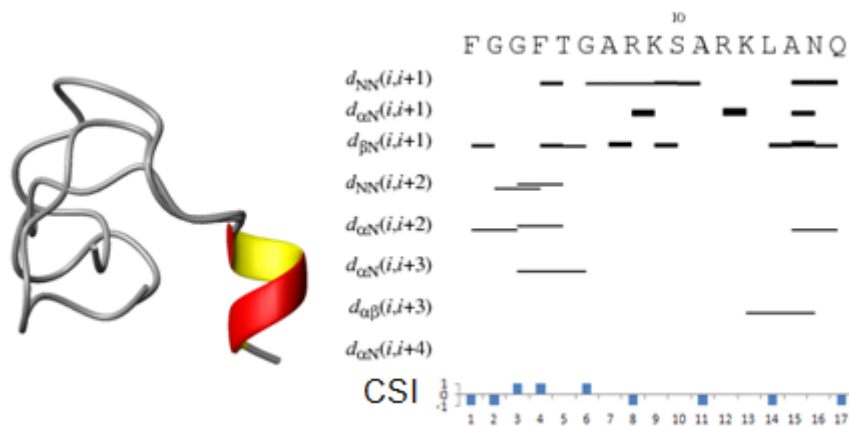


Figura 12: Estructuras CYANA del análogo [Ser¹⁰-O-β-D-GlcNAc]-nociceptina. Sólo 3 de las 20 estructuras calculadas presentan una α-hélice claramente definida.

En la figura 7 se observan las diferencias entre el NOESY de la nociceptina y su análogo [Thr⁵-O-α-D-GalNAc]-nociceptina. Estas diferencias señalan los cambios drásticos en la estructura de este análogo. El programa CYANA proporciona estructuras plegadas, que cumplen con los contactos remotos observados entre los residuos en el extremo carboxilo terminal y la región cercana a la glicosilación. La distribución de los ángulos de torsión del esqueleto de este análogo y de la nociceptina muestran los cambios estructurales entre ambos péptidos. En la nociceptina, la región más poblada

corresponde a una α -hélice. Sin embargo, en la [Thr⁵-O- α -D-GalNAc]-nociceptina la región mayoritaria corresponde a la estructura de PPII (Figura 13).

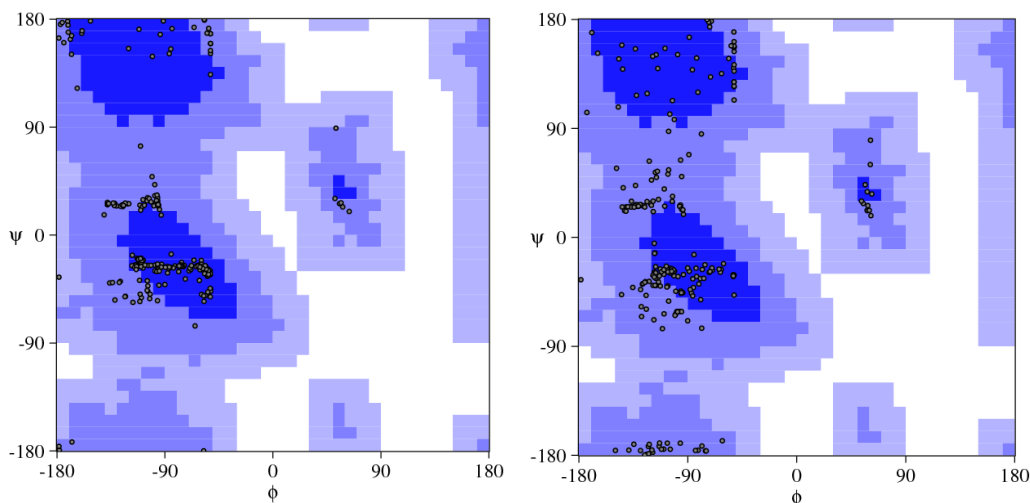


Figura 13: Mapas de Ramachandran de la nociceptina, *izquierda*, y del análogo [Thr⁵-O- α -D-GalNAc]-nociceptina, *derecha*. Se representan los valores de los ángulos ϕ y ψ de todos los aminoácidos de las 20 estructuras propuestas por CYANA.

Este análogo presenta una mayor rigidez en la región de glicosilación²²¹ pero el resto de la cadena muestra mayor flexibilidad que el resto de análogos. Aunque el número de contactos NOE interresiduales a tres y cuatro residuos desciende de manera sustancial, se sigue manteniendo una tendencia a α -hélice en la zona del carboxilo terminal como indican los CSI negativos (Figura 14). La rotación alrededor del enlace Thr- α -D-GalNAc está restringida. En la figura 14 se muestra la estructura minimizada de una geometría proporcionada por CYANA, que cumple todos los contactos NOE observados entre la α -D-GalNAc y la cadena de aminoácidos.

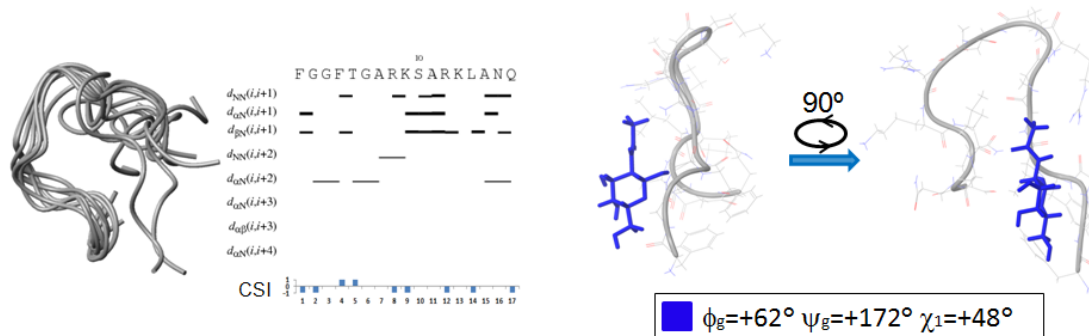


Figura 14: *Izda.*, estructuras de [Thr⁵-O- α -D-GalNAc]-nociceptina calculadas por CYANA, que cumplen los datos experimentales de NOESY. El diagrama muestra el resumen de los contactos NOE y los CSI. *Dcha.*, Estructura de menor energía calculada por CYANA con el residuo α -D-GalNAc unido a Thr⁵.

3.4. Conclusiones: Relación estructura-actividad

Las cuatro péptidos mostraron distintos valores de actividad frente a los receptores de membrana NOP (drNOP) desde el más activo, la Ser¹⁰-O- α -D-GalNAc]-nociceptina, a la Thr⁵-O- α -D-GalNAc]-nociceptina, con una afinidad un orden de magnitud inferior. Los cambios observados en su estructura podrían estar relacionados con la modulación de su actividad. Los datos estructurales muestran una mayor tendencia a formar α -hélice para la [Ser¹⁰-O- α -D-GalNAc]-nociceptina, coincidiendo con su mayor actividad biológica frente a los receptores NOP. Esta correlación entre el porcentaje de α -hélice y su actividad ha sido también propuesta en otros trabajos donde la tendencia de α -hélice de este péptido ha sido mejorada mediante restricciones covalentes entre distintos aminoácidos del extremo carboxilo terminal²²². La [Thr⁵-O- α -D-GalNAc]-nociceptina tiene menor tendencia a adoptar α -hélice, y muestra menor afinidad frente a estos receptores. Este glicopéptido adopta una estructura plegada que determina un descenso en la afinidad frente al receptor NOP.

El análisis estructural de estos análogos glicosilados ha puesto de manifiesto una relación estructura-función en la interacción entre la nociceptina y los receptores de membrana NOP. Estos resultados ilustran una estrategia en el diseño de péptidos análogos de nociceptina que potencien la estructura α -hélice en el extremo carboxilo terminal. La preorganización de este péptido en su extremo carboxilo terminal mejora la actividad frente a los receptores de membrana.

3.5. Materiales y métodos experimentales

3.5.1. Síntesis de péptidos

La nociceptina y sus 3 análogos glicosilados fueron producidos en el grupo Dr. Gregorio Valencia del IQAC-CSIC de Barcelona. Fueron sintetizados manualmente siguiendo una metodología estándar en fase sólida con resinas prederivatizadas de Wang¹⁷⁷, usando cuatro grupos protectores distintos para el acoplamiento. El paso limitante de la síntesis fue la producción de los aminoácidos glicosilados necesarios para ser utilizados como “building blocks” en la síntesis en fase sólida de los tres análogos glicosilados de la nociceptina²²³⁻²²⁵. En último lugar se liberó la cadena peptídica de la resina y se purificó con una columna de fase reversa de gradiente agua-acetonitrilo. Todos los péptidos fueron caracterizados mediante espectroscopía de masas de alta resolución (HRMS).

2.5.2. Cálculo de afinidades mediante ensayo competitivo frente a un radioligando

El cálculo de las afinidades frente a estos receptores de membrana NOP fue desarrollado por el grupo de la Dra. Raquel Rodríguez en el departamento de *Bioquímica y Biología Molecular* de la Facultad de Medicina de la Universidad de

Salamanca. Este ensayo de afinidad se realizó con la fracción de membrana de un cultivo de células HEK293 transfectadas, en las que se expresa el receptor de membrana NOP (*drNOP*) del pez cebra (*zebrafish*). Se midió la liberación de la sonda radiactiva [³H]-Nociceptina de la fracción de membranas por la adición de cantidades crecientes del péptido nociceptina no radiactivo y sus tres análogos glicosilados. La concentración del ligando radiactivo se mantuvo constante y la concentración de los péptidos objeto del estudio se varió desde 0.3 nM a 10 μM. La mezcla, fracción de membranas más el péptido radiactivo y el péptido analizado, tuvo un periodo previo de incubación de 1 hora a 298K. Tras varios pasos de filtración y lavado, la radiactividad fue medida mediante un contador de centelleo *Beckmann*.

3.5.3. Helicidad mediante dicroísmo circular

La medida de helicidad de estos péptidos se realizó por el grupo del Dr. Jesús T. Vázquez en el *Instituto Universitario de Bio-Organica Antonio González* de la Universidad de La Laguna. Los espectros de Dicroísmo Circular fueron adquiridos con un modelo de polarímetro JASCO J-600, en una cubeta de cuarzo cilíndrica de 1 mm. La elipticidad molar se calculó mediante la fórmula $[\theta]_r = \theta_{obs} \cdot MRW / (10 \cdot l \cdot [peptido])$, donde θ_{obs} es la elipticidad observada, *MRW* es la masa media por residuo (*Mean Residue Molecular Weight*) y *l* es el paso óptico.

3.5.4. Resonancia Magnética Nuclear

En primer lugar, se adquirieron espectros usando una mezcla 90:10 H₂O:D₂O a una temperatura de 278K en dos distintos campos magnéticos, espectrómetros *Bruker Avance 600* y *800 MHz*, con criosondas de triple canal. La asignación se abordó con experimentos 2D standard, TOCSY (distintos tiempos de mezcla [20, 60, 80 y 100 milisegundos]) y NOESY (dos tiempos de mezcla [200 y 300 milisegundos]). La concentración de las muestras de nociceptina y sus derivados glicosilados fue de 1-2 mM. Se controló el pH de las muestras para maximizar la observación de las señales que correlacionan el resto de señales con los hidrogenos intercambiables de las amidas. El pH de la nociceptina fue ajustado a 5,4 y el de sus tres análogos glicosilados a 6,7.

En segundo lugar se adquirieron los espectros en presencia de micelas de SDS (*Sodium Dodecyl Sulfate*). Este agente mimetiza el medio en las regiones cercanas a una membrana celular. Se obtuvieron la misma serie de experimentos que en ausencia de SDS, a una temperatura de 288K en un espectrómetro *Bruker Avance 600 MHz*. Los péptidos se analizaron a una concentración de 2mM en presencia de SDS deuterado a una concentración de 132mM. La señal del TSP (*2,2,3,3-tetradeutero-3-TrimethylSilylPropionic acid*) fue utilizada como referencia interna de los desplazamientos químicos de protón ($\delta TSP = 0$ ppm).

Nociceptina				
a.a.	NH	H α	H β	misc.
Phe ¹		4,512	3,091 ; 2,937	7,114 (H δ); 7,006 (H ϵ); 6,706 (H ζ)
Gly ²	8,232	3,507 ; 3,604		
Gly ³	8,186	3,981 ; 3,742		
Phe ⁴	7,864	4,564	3,236 ; 2,846	7,241 (H δ); 7,090 (H ϵ); 6,952(H ζ)
Thr ⁵	8,320	4,169	4,101	1,029(H γ)
Gly ⁶	8,341	4,162		
Ala ⁷	7,836	3,928	1,220	
Arg ⁸	8,176	4,147	1,683	1,536(H γ); 2,965(H δ); 7,049(H ϵ)
Lys ⁹	7,769	3,991	1,651	1,261(H γ); 1,459(H δ); 2,797(H ϵ)
Ser ¹⁰	7,930	4,065	3,681	
Ala ¹¹	7,745	4,003	1,258	
Arg ¹²	7,704	3,944	1,675	1,508(H γ); 2,976(H δ); 7,031(H ϵ)
Lys ¹³	7,783	3,809	1,683	1,242(H γ); 1,506(H δ); 2,799(H ϵ)
Leu ¹⁴	7,744	3,999	1,499	1,378(H γ); 0,687(H δ)
Ala ¹⁵	7,665	4,052	1,222	
Asn ¹⁶	7,851	4,545	2,681 ; 2,523	7,428(H δ 21); 6,727(H δ 22)
Gln ¹⁷	7,561	3,929	1,934 ; 1,731	2,129(H γ); 7,357(H ϵ 21); 6,658(H ϵ 22)

(Ser ¹⁰ -O- α -D-GalNAc)-Nociceptina									
a.a.	NH	H α	H β	misc.					
Phe ¹		4,164	2,968 ; 3,104	7,124(H δ); 7,021(H ϵ); 6,708(H ζ)					
Gly ²	8,232	3,511 ; 3,603							
Gly ³	8,241	4,004 ; 3,747							
Phe ⁴	7,851	4,595	2,846 ; 3,275	7,269(H δ); 7,111(H ϵ); 6,955 (H ζ)					
Thr ⁵	8,414	4,201		1,042(H γ)					
Gly ⁶	8,390	4,202							
Ala ⁷	7,786	3,965	1,227						
Arg ⁸	8,198	3,958	1,692	1,514(H γ); 2,969(H δ); 7,074(H ϵ)					
Lys ⁹	7,767	3,789	1,650	1,256(H γ); 1,520(H δ); 2,801(H ϵ)					
Ser ¹⁰	7,859	4,271	3,837						
Ala ¹¹	7,910	3,939	1,287						
Arg ¹²	7,894	3,977	1,700	1,492(H γ); 2,989(H δ); 7,029(H ϵ)					
Lys ¹³	7,775	4,002	1,669	1,292(H γ); 1,536(H δ); 2,804(H ϵ)					
Leu ¹⁴	7,600	4,015	1,491	1,391(H γ); 0,708(H δ)					
Ala ¹⁵	7,661	4,053	1,242						
Asn ¹⁶	7,823	4,562	2,699 ; 2,532	7,411(H δ 21); 6,736 (H δ 22)					
Gln ¹⁷	7,543	3,930	1,947 ; 1,756	2,158(H γ); 7,368(H ϵ 21); 6,669(H ϵ 22)					
	H1	H2	H3	H4	H5	H6 ₁	H6 ₂	NH	CH ₃
α -D-GalNAc	4,662	3,982	3,632	3,774		3,691	3,426	8,049	1,823

Figura 15: Los desplazamientos químicos de la nociceptina (*arriba*) y del glicopéptido [Ser¹⁰-O- α -D-GalNAc]-nociceptina (*abajo*) en presencia de SDS. Los valores están referenciados al singlete de la referencia interna TSP (0ppm).

(Ser ¹⁰ -O-β-D-GlcNAc)-Nociceptina									
a.a.	NH		Hα	Hβ		misc.			
Phe ¹			4,159	2,956 ; 3,088		7,124(Hδ); 7,023(Hε); 6,734(Hζ)			
Gly ²	8,225		3,577						
Gly ³	8,141		3,971 ; 3,701						
Phe ⁴	7,862		4,571	3,226 ; 2,858		7,242(Hδ); 7,100(Hε); 6,957(Hζ)			
Thr ⁵	8,265		4,159	4,070		1,035(Hγ)			
Gly ⁶	8,264		4,149						
Ala ⁷	7,805		4,083	1,233					
Arg ⁸	8,141		3,984	1,676		1,486(Hγ); 2,995(Hδ); 7,030(Hε)			
Lys ⁹	7,792		4,060	1,637		1,304(Hγ); 1,518(Hδ); 2,808 (Hε)			
Ser ¹⁰	8,004		4,187	3,925 ; 3,711					
Ala ¹¹	7,788		3,988	1,227					
Arg ¹²	7,761		4,021	1,653		1,501(Hγ); 2,990(Hδ); 6,974(Hε)			
Lys ¹³	7,750		4,016	1,671		1,268(Hγ); 1,530(Hδ); 2,795(Hε)			
Leu ¹⁴	7,740		4,008	1,516		1,390(Hγ); 0,701(Hδ)			
Ala ¹⁵	7,675		4,073	1,218					
Asn ¹⁶	7,905		4,534	2,664 ; 2,524		7,415(Hδ21); 6,724(Hδ22)			
Gln ¹⁷	7,600		3,923	1,939 ; 1,746		2,117(Hγ); 7,350(Hε21); 6,653(Hε22)			
	H1	H2	H3	H4	H5	H6 ₁	H6 ₂	NH	CH ₃
β-D-GlcNAc	4,356	3,483	3,348	3,241	3,232	3,726	3,404	7,955	1,805

(Thr ⁵ -O-α-D-GalNAc)-Nociceptina									
a.a.	NH	Hα	Hβ	misc,					
Phe ¹		4,153	2,949 ; 3,091	7,129(Hδ); 6,759(Hε); 7,039(Hζ)					
Gly ²	8,212	3,590							
Gly ³									
Phe ⁴	7,909	4,606	2,927 ; 3,188	7,249(Hδ); 7,112(Hε); 6,977(Hζ)					
Thr ⁵	8,240	4,401	4,269	0,969(Hγ)					
Gly ⁶	8,240*	4,401*							
Ala ⁷	8,163	4,121	1,216						
Arg ⁸	8,085	4,002	1,682	1,491(Hγ); 2,983(Hδ); 7,008(Hε)					
Lys ⁹	7,879	3,972	1,630	1,241(Hγ); 1,508(Hδ); 2,803(Hε)					
Ser ¹⁰	7,839	4,163	3,740 ; 3,656						
Ala ¹¹	7,932	4,025	1,241						
Arg ¹²	7,783	3,975	1,655	1,500(Hγ); 3,015(Hδ); 7,000(Hε)					
Lys ¹³	7,800	4,065	1,642	1,255(Hγ); 1,504(Hδ); 2,801(Hε)					
Leu ¹⁴	7,770	4,010	1,513	1,376(Hγ); 0,699(Hδ)					
Ala ¹⁵	7,704	4,057	1,214						
Asn ¹⁶	7,921	4,534	2,677 ; 2,523	7,430(Hδ21); 6,725(Hδ22)					
Gln ¹⁷	7,623	3,923	1,942 ; 1,736	2,122(Hγ); 6,655(Hε21); 7,350(Hε22)					
*solapamiento									
	H1	H2	H3	H4	H5	H6 ₁	H6 ₂	NH	CH ₃
α-D-GalNAc	4,686	3,791	3,528			3,694	3,451	7,717	1,835

Figura 16: Los desplazamientos químicos del glicopéptido [Ser¹⁰-O-β-D-GlcNAc]-nociceptina (*arriba*) y del glicopéptido [Thr⁵-O-α-D-GalNAc]-nociceptina (*abajo*) en presencia de SDS. Los valores están referenciados al singlete de la referencia interna TSP (0ppm).

3.5.5. Estructura tridimensional

El cálculo de las estructuras de estos cuatro péptidos se realizó a partir de la lista de picos de los espectro NOESY que se adquirieron en la muestras en presencia de micelas de SDS. En primer lugar, se llevó a cabo la asignación de los sistemas de spin de los aminoácidos mediante experimentos TOCSY y de los picos de cruce secuenciales de los espectros NOESY usando el software CARA¹³². Seguidamente, se seleccionaron todos los picos de cruce de los espectros NOESY mediante CARA, obteniendo una lista de picos de cruce y de sus volúmenes correspondientes. Esta lista de volúmenes se usó como archivo de entrada del programa CYANA (version 2.1)^{131,226}, que genera, por cálculos de dinámica molecular restringida, estructuras compatibles con los datos de NOE experimentales mediante variaciones de los grados de libertad de torsión. Este cálculo se realizó teniendo en cuenta únicamente la cadena de aminoácidos, sin el componente glicosídico. Esta parte de la molécula se añadió posteriormente, de manera manual, teniendo en cuenta los datos experimentales de los espectros NOESY. Un cálculo de CYANA consta de siete ciclos que combinan la asignación de los picos de cruce del NOESY, según la lista de desplazamientos químicos introducida, y el cálculo de estructuras, seguidos de un último cálculo de la estructura del péptido. Cada uno de los cálculos estructurales viene precedido por la generación estocástica por variación de los ángulos torsionales de 100 conformaciones. Tras cada uno de los ciclos, las 20 conformaciones de menor valor CYANA de confluencia (menor desviación de los datos experimentales) son utilizadas para el siguiente ciclo. Tras el último cálculo estructural, el programa genera 20 conformaciones que representan la contribución de todos los picos de cruce del espectro NOESY según la asignación inicial. Además, CYANA genera el mapa de Ramachandran de todos los aminoácidos de estas 20 conformaciones y el valor RMSD (Root Mean Square Deviation) para el esqueleto peptídico y las cadenas laterales de cada uno de los conformeros frente a la estructura promedio. Las 20 estructuras tridimensionales obtenidas del cálculo de CYANA fueron representadas mediante el programa de visualización MOLMOL²²⁷.

	PÉPTIDOS			
	Nociceptina	(Ser ¹⁰ -O- α -D-GalNAc)- Nociceptina	(Ser ¹⁰ -O- β -D-GlcNAc)- Nociceptina	(Thr ⁵ -O- α -D-GalNAc)- Nociceptina
NOE upper distant limits				
Short-range, $ i-j \leq 1$	162	143	139	157
Medium-range, $1 < i-j < 5$	36	45	34	20
Long-range, $ i-j \geq 5$	6	0	8	7
CYANA target function (\AA^2)	0.01	0.01	0.03	0.01
RMSD to mean coordinates (\AA)				
Backbone	1.69	2.68	2.03	2.76
All heavy atoms	2.11	3.59	2.87	3.47
Ramachandran plot statistics (%)				
Residues in the most favoured	59.6	70	57.9	44.6
Residues in the additionally	40	30	42.1	55
Residues in the generously	0.4	0	0	0.4
Residues in the disallowed	0	0	0	0

Tabla 1: Estadísticas de las estructuras de los cuatro péptidos en presencia de micelas de SDS. Los valores del programa CYANA hacen referencia a los 20 conformeros del último cálculo estructural.

4. El uso de etiquetas difluoroacetamida como sonda para detectar interacciones carbohidrato-proteína

Las lectinas son proteínas capaces de reconocer carbohidratos y que no forman parte de la superfamilia de las inmunoglobulinas, ni son enzimas²²⁸. Fueron descubiertas en el siglo XIX por su capacidad aglutinante de eritrocitos. Su estudio aumentó a partir de la segunda mitad del siglo XX cuando se encontró que muchas de ellas eran específicas para determinados tipos de azúcares²²⁹. Estas proteínas se pueden encontrar en todo tipo de organismos y están implicadas en el reconocimiento molecular de glicanos de diversa naturaleza²³⁰.

Una de las lectinas más estudiadas ha sido la aglutinina de germen de trigo (*Wheat Germ Agglutinin*, WGA). Se extrae del trigo (*Triticum vulgare* o *Triticum aestivum*), y se purifica por cromatografía de afinidad²³¹. Como su propio nombre indica, esta lectina tiene la capacidad de aglutinar los glóbulos rojos mediante el reconocimiento de los antígenos de la superficie de estas células sanguíneas, pero también ha sido utilizada para discernir entre células cancerosas y células sanas²³²⁻²³³. Asimismo, esta lectina inmovilizada sobre soporte sólido se utiliza de forma habitual como reactivo para la identificación de proteínas glicosiladas²³⁴⁻²³⁵.

En la extracción de esta lectina se obtienen tres isoformas, WGA1 (35%), WGA2 (55%) y WGA3 (10%), que son muy parecidas entre sí²³⁶. Estas isoformas son proteínas homodiméricas de un peso molecular de 36 kDa. Cada uno de sus protómeros contiene cuatro dominios de tipo heveína²³⁷⁻²³⁸. Los correspondientes sitios de unión se identifican según los dominios de heveína implicados (A, B, C y D), y la cadena polipeptídica a la que pertenecen (1 y 2). Siguiendo esta terminología, uno de estos homodímeros contiene 8 sitios de reconocimiento que se denominan A1, B1C2, C1B2, D1A2, A2, B2C1, C2B1 y D2A1 (Figura 1)²³⁹. Aquellos sitios en los que están implicados los dominios A y D presentan menor afinidad²⁴⁰.

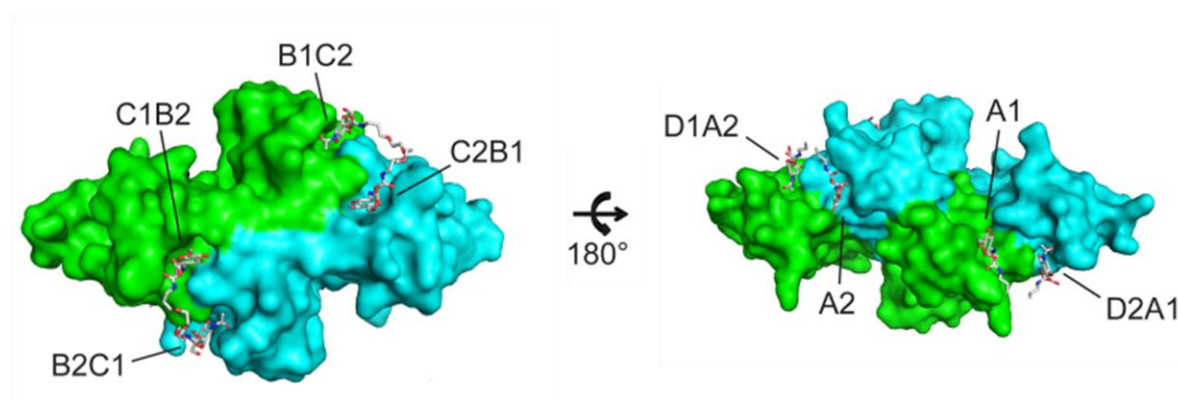


Figura 1: Estructura tridimensional del homodímero de la lectina WGA (según PDB: 2X52). En esta representación se representa el protómero 1 en color verde y el protómero 2 en color cian, donde se identifican los ocho centros de reconocimiento de la lectina²³⁹.

La estructura e interacciones de las tres isoformas de la WGA con diversos ligandos han sido caracterizadas en detalle por estudios de cristalografía de Rayos X^{239,241-244} y de Resonancia Magnética Nuclear^{236,245}. Se han estudiado complejos con glicanos cuyo extremo no-reductor es ácido siálico, así como otros que muestran N-acetil-glucosamina (GlcNAc). El reconocimiento molecular está mediado por aminoácidos aromáticos, que tienen interacciones de apilamiento (*stacking*) con el anillo de piranosa, y por aminoácidos polares, que forman enlaces de hidrógeno. Por ejemplo, el sitio de unión C1B2 de *WGA1* interactúa con el residuo de GlcNAc mediante *stacking* con una fenilalanina de la cadena 1, Phe¹⁰⁹, y se complementa con una red de enlaces de hidrógeno con los residuos Asp⁸⁶ y Ser¹⁰⁵ de la cadena 1 y de los residuos Ala⁷¹ y Glu⁷² de la cadena 2 (Figura 2, *izda.*). En el caso de *WGA3*, los sitios de unión C2B1 o C1B2 muestran un patrón de interacciones similares a las descritas para *WGA1*. La isoforma *WGA3* interactúa con un residuo de GlcNAc a través de Tyr¹⁰⁹, además de presentar los mismos enlaces de hidrógeno con los residuos Asp⁸⁶, Ser¹⁰⁵ y Glu⁷². Otros sitios de unión como B1C2 y B2C1 muestran *stacking* a través de su His⁶⁶ y su red de enlaces de hidrógeno implica a los aminoácidos Ser⁴³, Ser⁶², Tyr⁷³, Ser¹¹⁴ y Glu¹¹⁵ (Figura 2, *dcha.*). El modo de interacción de estos ligandos que contienen residuos GlcNAc se ha observado también en complejos de estas isoformas con ligandos que presentan ácido siálico como extremo no reductor²⁴¹.

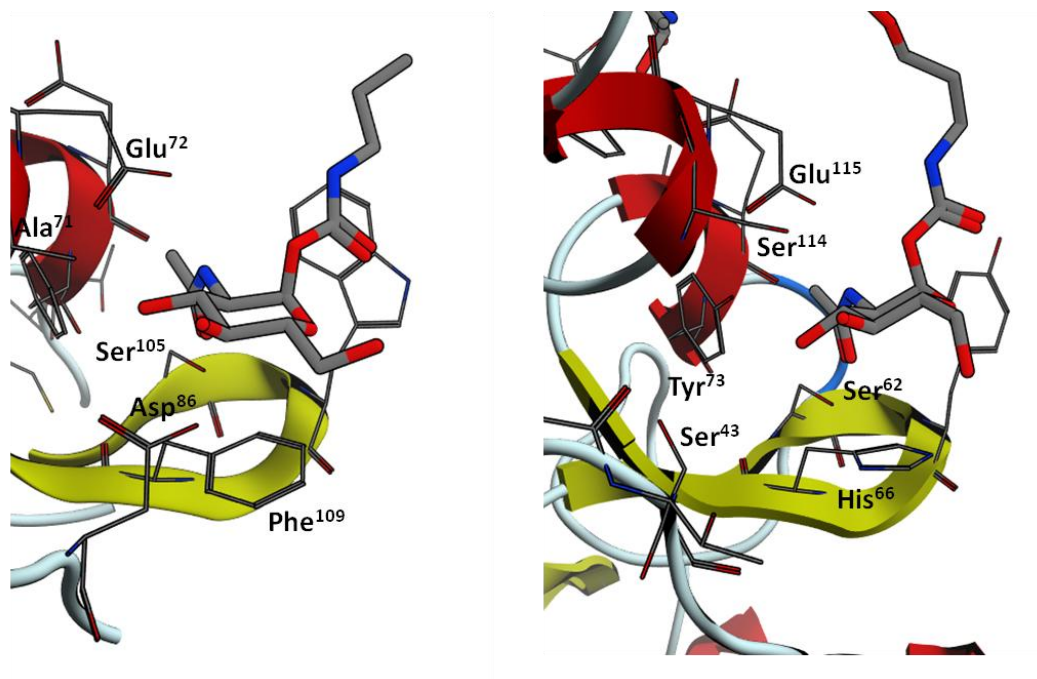


Figura 2: Sitios de unión en distintas isoformas de WGA. *Izda.*, detalle del sitio de unión C1B2 de la isoforma *WGA1* (PDB: 2UWG) *Dcha.*, detalle del sitio de unión B1C2 de la isoforma *WGA3* ((PDB: 2X52). Los residuos implicados en el reconocimiento de la GlcNAc están numerados.

Además de caracterizarse la interacción de esta lectina con distintos oligosacáridos que contienen dos o más residuos de GlcNAc²⁴⁶ o de ácido siálico²⁴⁷ también se ha encontrado que es capaz de unir otros oligosacáridos con al menos un residuo de NAcetil-glucosamina como por ejemplo, GlcNAc- β -(1-6)-1-O-Me- α -D-Man²⁴⁸, GlcNAc-(1-6)-D-Gal²⁴⁴, o GlcNAc-(1-3)- β -D-Gal-(1-4)-D-Glc²⁴¹. De hecho, estos son algunos de los epítomos presentes en las cadenas glucídicas de diversas glicoproteínas presentes en la superficie celular²⁴⁹⁻²⁵⁰. Además, la presencia de múltiples sitios de unión tanto en el monómero como en la propia estructura dimérica facilita a la WGA su participación en procesos de interacción multivalente. La estructura dimérica de WGA podrá interaccionar simultáneamente a través de varios de sus sitios de unión con un ligando que contenga varios epítomos separados, al menos, por la distancia mínima entre dos de sus sitios de unión de WGA. Wittmann y col. han sintetizado ligandos capaces de establecer uniones simultaneas mediante los residuos GlcNAc que contienen en sus extremos, y han podido determinar experimentalmente que la distancia mínima entre dos sitios de reconocimiento de WGA es de 13-14 Å^{239,251-254}.

En este capítulo se aborda el estudio de la interacción de análogos fluorados de la N-acetil-glucosamina frente a la WGA, como lectina modelo. La introducción de motivos fluorados en el diseño de ligandos se utiliza frecuentemente en química farmacéutica, ya que puede mejorar las propiedades farmacocinéticas y fisicoquímicas de los compuestos. De hecho, actualmente el 20% de los medicamentos comercializados contiene al menos un átomo de flúor²⁵⁵⁻²⁵⁸.

En esta tesis se han estudiado derivados de GlcNAc donde el resto acetilo ha sido substituido por grupos difluoroacetilo que han sido sintetizados en el laboratorio del Dr. Niels Reichardt, en CIC biomaGUNE, mediante una reacción directa de acilación en la que se han añadido los fragmentos de difluoroacetilo a los aminoazúcares. La presencia de átomos de flúor en el fragmento de difluoroacetamida debería permitir seguir fácilmente mediante RMN la posible interacción del glicano modificado con el receptor correspondiente. En este contexto, en este capítulo se muestra que distintos glicanos que contienen esta etiqueta fluorada mantienen las propiedades de interacción de sus análogos naturales cuando interaccionan con la lectina modelo WGA, mientras que la interacción CH- π del motivo difluoroacetamida es incluso más intensa que con el grupo acetamido o el grupo trifluoroacetamido.

4.1. Análogos fluorados de quitobiosa y quitotriosa en estado libre

La asignación de las señales de RMN de ¹H y ¹⁹F de estas moléculas es necesaria para identificar posteriormente la posición de los residuos que son afectados en presencia de la proteína. La repetición de dos o tres unidades de GlcNAc (Figura 3) provoca la existencia de un alto solapamiento de sus señales de RMN.

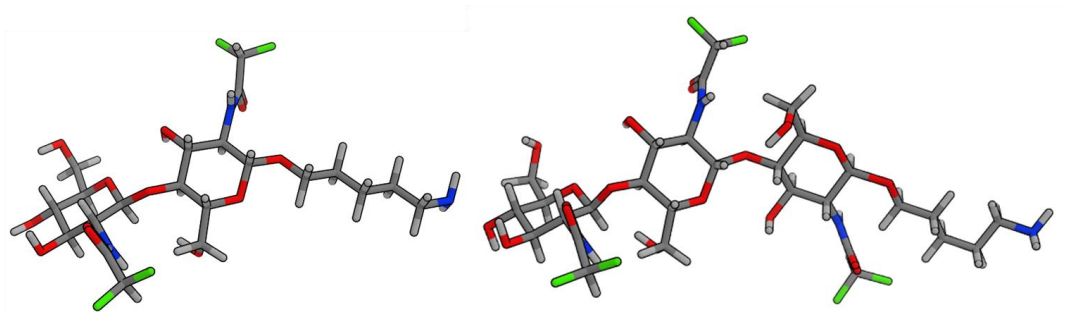


Figura 3: Estructura de los análogos de N,N'-diacetil quitobiosa (*Izda.*) y N,N',N''-triacetil quitotriosa (*Dcha.*) con etiquetas difluoroacetamida. El grupo acetilo de estos ligandos contiene dos átomos de flúor y un protón, $-\text{COCHF}_2$. El extremo reductor está protegido por una cadena hidrocarbonada de pentil amina.

El ligando de quitobiosa mostró un elevado solapamiento en el espectro de ^1H RMN (Figura 4). La únicas señales de los residuos de GlcNAc que están aisladas son las de los protones anoméricos, H1, los protones amídicos y los protones del grupo difluoroacetamido (Figuras 4 y 5). Los protones del grupo difluoroacetamido son fácilmente reconocibles por su aspecto de triplete entre 6 y 6.5 ppm debido al fuerte acoplamiento heteronuclear con los núcleos de ^{19}F (Figura 5). Esta zona está libre de otras señales, por lo que es una zona excelente para seguir la posible interacción. A pesar del solapamiento, la asignación de los distintos sistemas de espín (Figuras 4, 5 y 6) se consiguió mediante experimentos homonucleares estándar 2D COSY, TOCSY y NOESY.

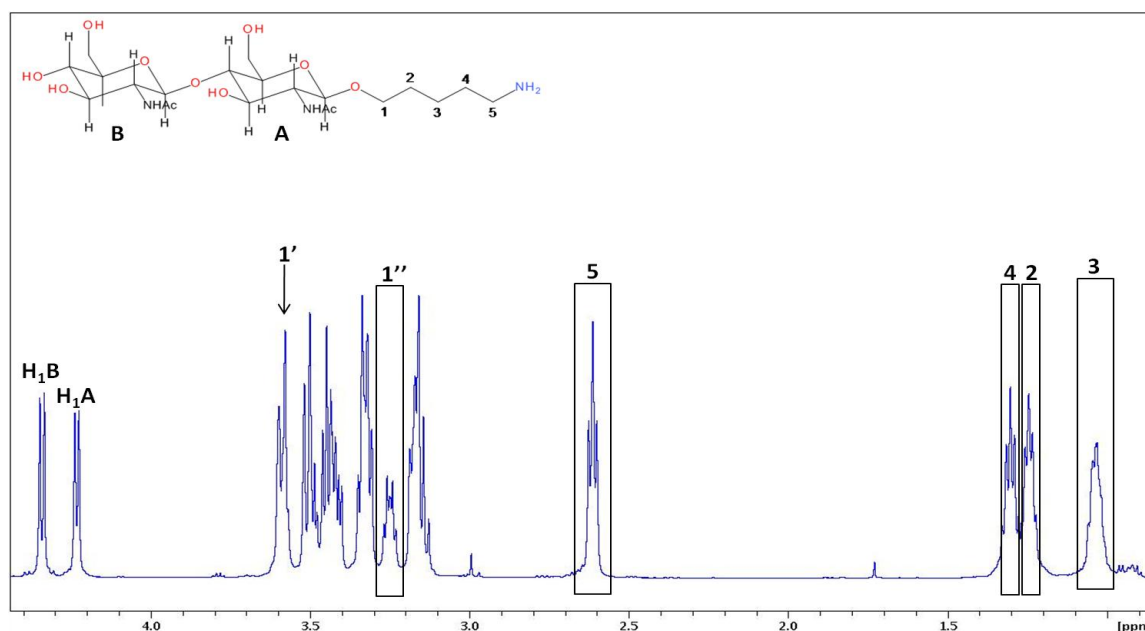


Figura 4: Zona campo alto (0.8-4,5 ppm) del espectro 1D ^1H RMN realizado en D_2O (278K) del derivado de N,N'-diacetil quitobiosa con etiquetas difluoroacetamida. Los protones anoméricos de los residuos de GlcNAc, H₁A y H₁B se diferencian fácilmente del resto de protones. En esta figura se identifican los grupos metileno de la cadena de pentil amina.

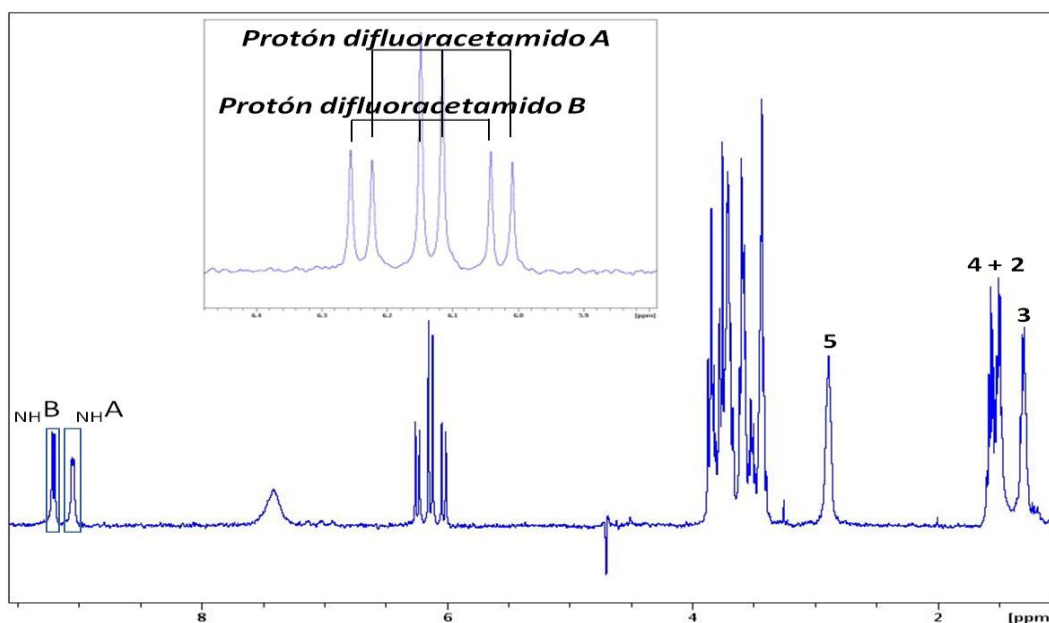


Figura 5: Espectro completo 1D ^1H RMN realizado en una mezcla 90:10 $\text{H}_2\text{O}:\text{D}_2\text{O}$ (tampón fosfato 50mM a pH 6.0, 298K) y con supresión de la señal de agua (las señales de los protones anómicos entre 4-4,5 ppm quedan también suprimidas) del derivado de N,N'-diacetil quitobiosa con etiquetas difluoroacetamida. Los protones de estos grupos son fácilmente reconocibles por el desdoblamiento de sus señales en forma de tripletes por el acoplamiento que presentan con los dos átomos de flúor.

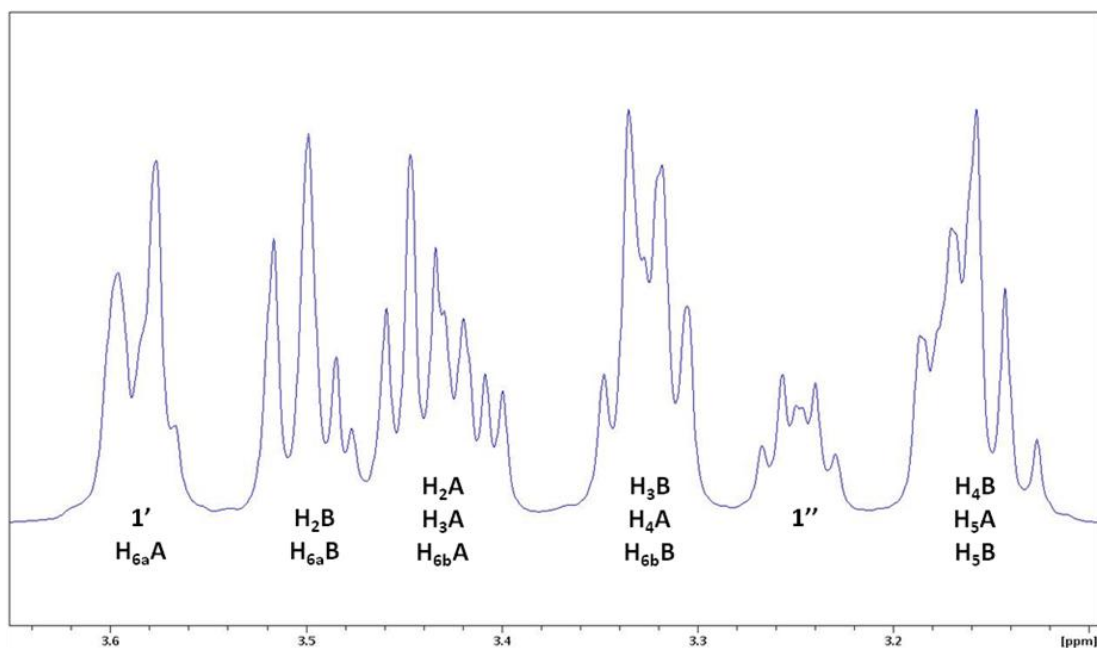


Figura 6: Detalle del espectro ^1H RMN del derivado de N,N'-diacetil quitobiosa con etiquetas difluoroacetamida, mostrando el solapamiento existente. La asignación de los protones de los residuos de GlcNAc se especifica bajo los multipletes.

Los átomos de flúor de estos residuos son diastereotópicos y producen dos sistemas AA'BB' fuertemente acoplados. De hecho, sus diferencias de desplazamiento químico (20-40 Hz) son menores que la constante homonuclear entre ellos ($^2J_{FF} \sim 300\text{Hz}$) dando lugar a acoplamientos de segundo orden, que se manifiestan presentando un fuerte *efecto tejado* que maximiza uno de los picos del doblete y minimiza el otro. La proyección del espectro 2D ^1H - ^{19}F de la figura 7 muestra los picos de baja intensidad. Obviamente, no se debe confundir esta constante homonuclear ^{19}F - ^{19}F con la constante heteronuclear ^1H - ^{19}F ($^2J_{HF} \sim 54\text{Hz}$) que se observa en las señales de ^{19}F en los experimentos convencionales sin desacoplamiento de ^1H (Figura 7, *expansión a la derecha*).

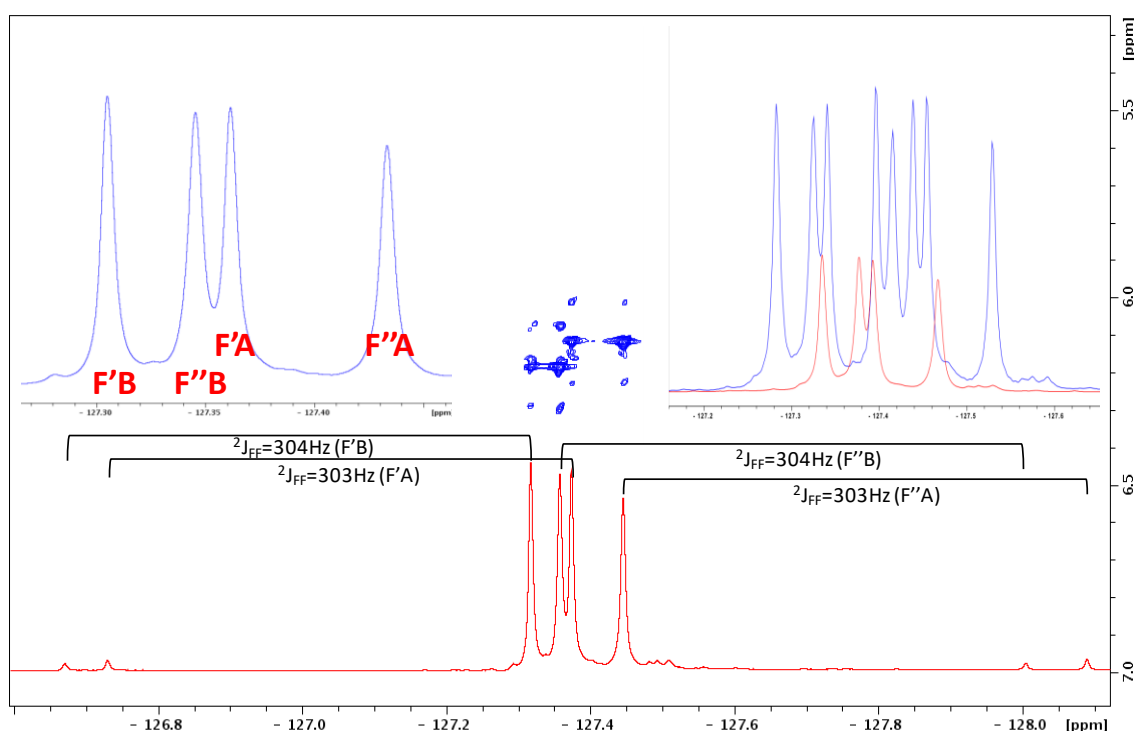


Figura 7: Espectro de correlación heteronuclear ^{19}F - ^1H del derivado de N,N'-diacetil quitobiosa con etiquetas difluoroacetamida. Este experimento permite correlacionar los protones del motivo difluoroacetamida con sus respectivos ^{19}F . La *inserción a la izquierda* muestra la asignación de las señales de flúor de los motivos GlcNAc. En la *inserción a la derecha* se superpone un espectro monodimensional de flúor sin desacoplamiento de protón (*espectro azul*) y un espectro de flúor con desacoplamiento de ^1H durante la adquisición (*espectro rojo*).

Siguiendo un procedimiento similar, se asignaron los espectros de ^1H y ^{19}F del análogo de quitotriosa. El solapamiento en este compuesto es todavía mayor que en el disacárido. De nuevo, los protones H1 (Figura 8), los amídicos y los del grupo difluoroacetamido (Figura 9) están aislados del resto, aunque dos de ellos solapan a 500 MHz (residuos B y C). La asignación de los protones amídicos se realizó mediante una serie de espectros NOE selectivos (Figura 10). A pesar del solapamiento, se pudieron asignar todas las señales de este compuesto mediante experimentos 2D convencionales (Figura 11).

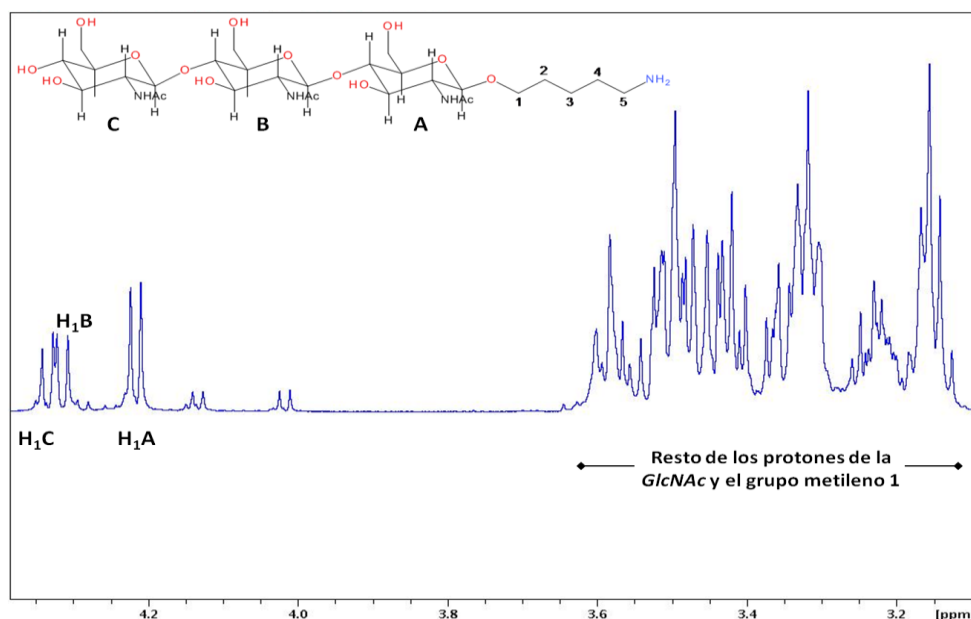


Figura 8: Espectro 1D ^1H RMN realizado en una mezcla 90:10 $\text{H}_2\text{O}:\text{D}_2\text{O}$ (tampón fosfato 50mM a pH 6.0, 278K) del derivado de N,N',N''-triacetil quitotriosa con etiquetas difluoroacetamida. Los protones H1 de los residuos de GlcNAc se diferencian fácilmente del resto de protones del anillo de piranosa. Entre 4-4,2 ppm se observan las trazas de un contaminante que no ha podido ser eliminadas en el proceso de purificación de esta muestra.

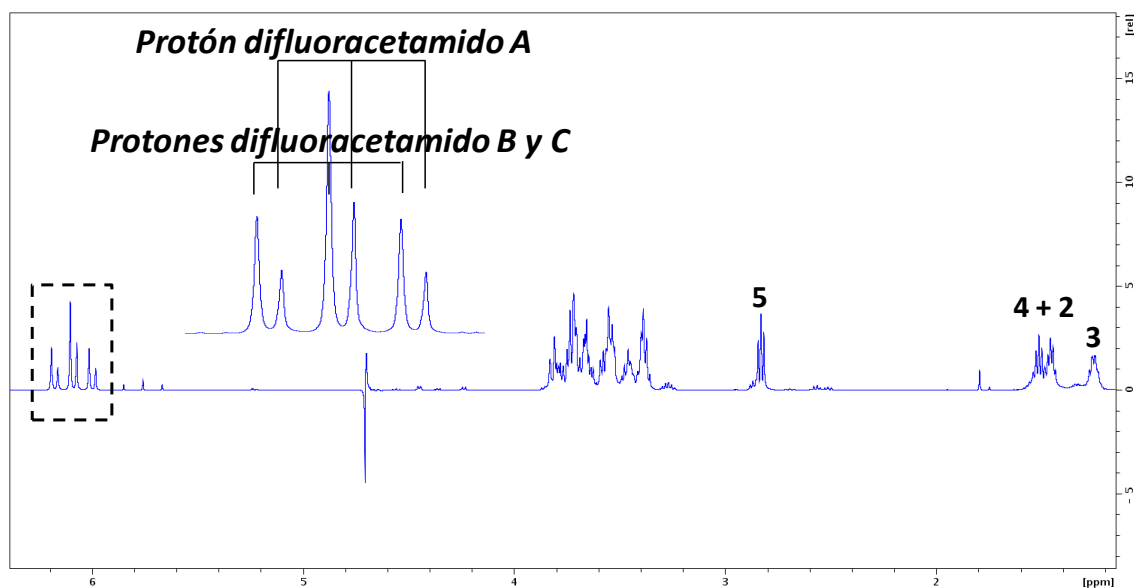


Figura 9: Espectro 1D ^1H RMN realizado en una mezcla 90:10 $\text{H}_2\text{O}:\text{D}_2\text{O}$ (298K) y con supresión de la señal de agua (las señales de los protones anómicos entre 4,2-4,7 ppm quedan también suprimidas) del derivado de N,N',N''-triacetyl quitotriosa con etiquetas difluoroacetamida. *Inserción*, Los protones difluoroacetamido son fácilmente reconocibles por el desdoblamiento de sus señales en forma de tripletes por el fuerte acoplamiento que presentan con los dos átomos de flúor. Las señales de los protones del grupo difluoroacetamido de los residuos B y C están completamente solapadas.

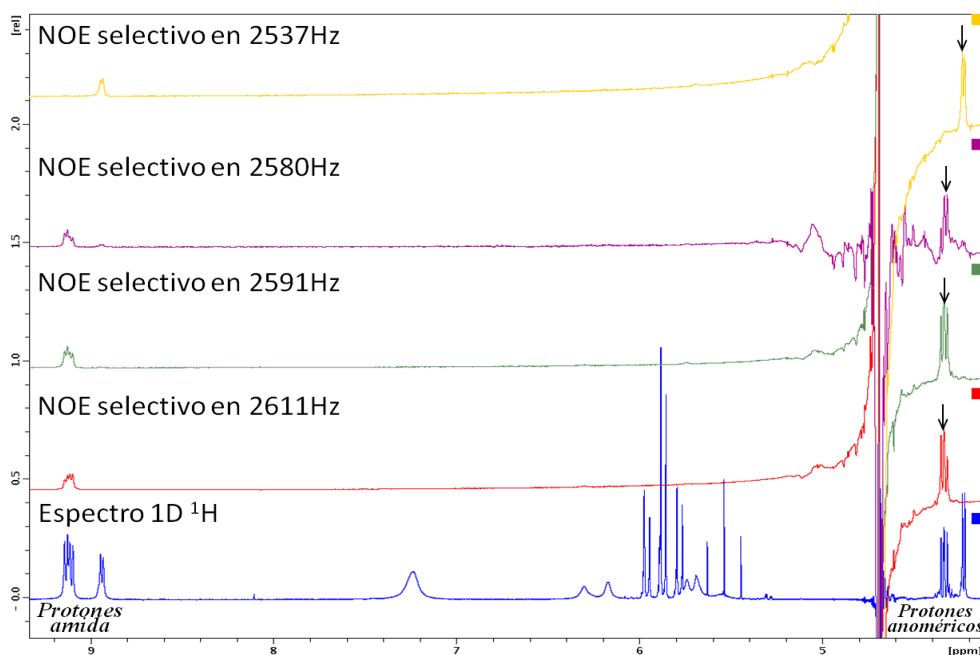


Figura 10: Espectro 1D ^1H RMN en una mezcla 90:10 $\text{H}_2\text{O}:\text{D}_2\text{O}$ (278K) y espectros de NOE selectivo del derivado de N,N',N''-triacetil quitotriosa con etiquetas difluoroacetamida. Los contactos NOE entre los protones H1 y los NH de los residuos de GlcNAc modificados permitieron la asignación de los protones amida de cada residuo. Entre 5,4-5,6 ppm se observan trazas de difluoroacético que no ha podido ser eliminado en el proceso de purificación de esta muestra.

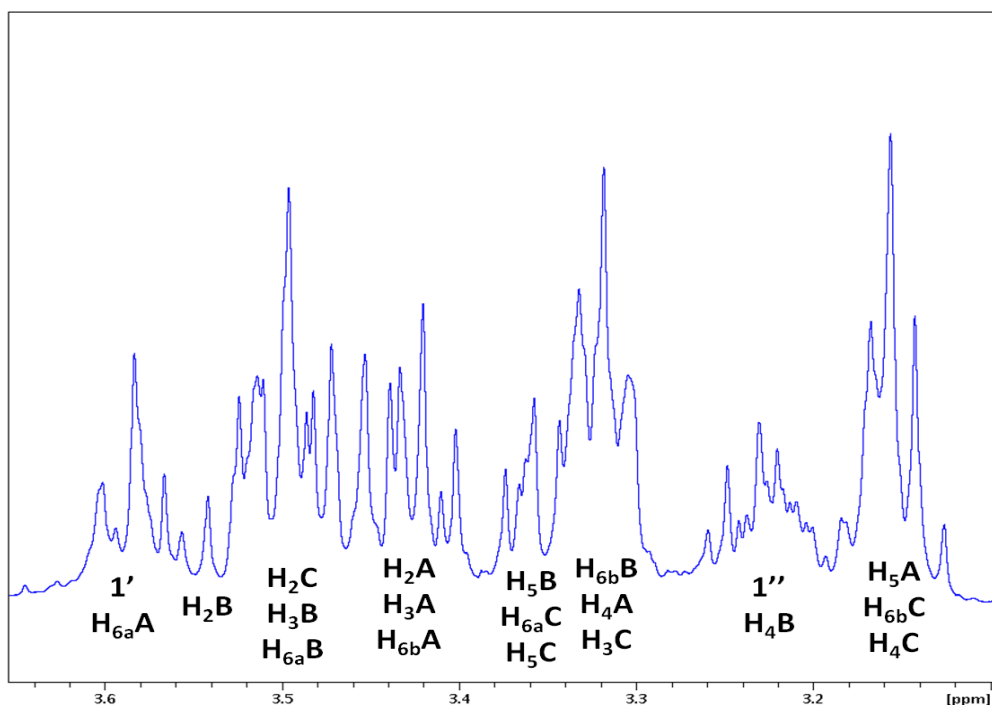


Figura 11: Detalle del espectro 1D ^1H RMN del derivado de N,N',N''-triacetil quitotriosa con etiquetas difluoroacetamida. La asignación de los protones de los residuos de GlcNAc se especifica bajo los multipletes.

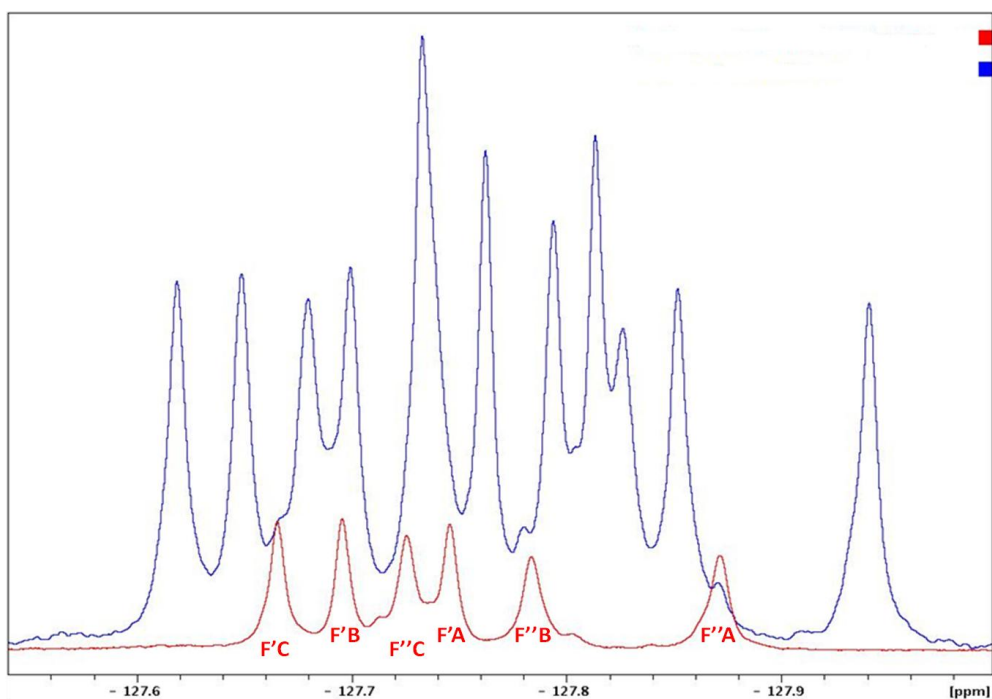


Figura 12 Espectros de 1D ^{19}F RMN del derivado de N,N',N''-triacetil quitotriosa con etiquetas difluoroacetamida. Se superponen dos espectros, sin desacoplamiento de ^1H (azul) y con desacoplamiento (rojo) durante la adquisición.

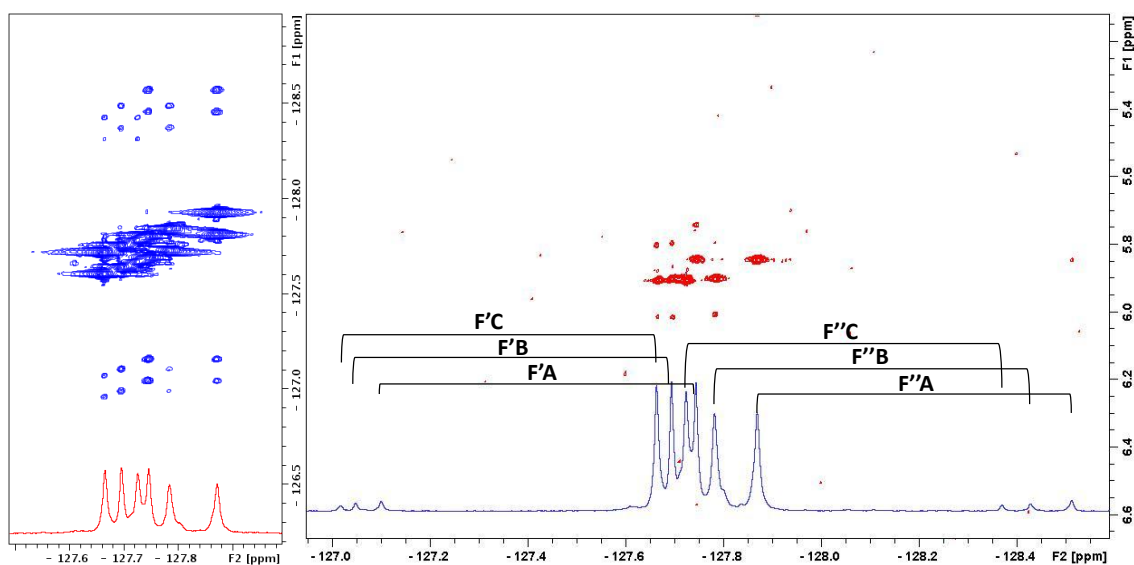


Figura 13: *Izda.*, espectro 2D ^{19}F , ^{19}F -COSY del derivado de N,N',N''-triacetil quitotriosa con etiquetas difluoroacetamida. La proyección en *rojo* es un espectro 1D ^{19}F RMN de las señales intensas de los tres sistemas AA'BB'. *Dcha.*, espectro 2D de correlación heteronuclear ^1H - ^{19}F del derivado de N,N',N''-triacetil quitotriosa con etiquetas difluoroacetamida. La proyección en *azul* muestra un espectro completo 1D ^{19}F RMN en el que se han destacado las parejas de picos de cada uno de los dobletes de los tres sistemas AA'BB'.

Las señales de 1D ^{19}F de nuevo mostraron un fuerte efecto *tejado*. Existen 6 señales muy intensas y otras seis muy débiles, separadas por la constante de acoplamiento homonuclear, $^2J_{\text{FF}} \sim 300\text{Hz}$ (Figura 13). El espectro homonuclear 2D ^{19}F , ^{19}F -COSY permitió correlacionar los núcleos de flúor de cada grupo difluoroacetamida (Figura 13 *izquierda*), identificando que señales de flúor pertenecen al mismo residuo GlcNAc. Por otro lado, el experimento heteronuclear permitió distinguir únicamente que señales de flúor pertenecen al extremo reductor, residuo A (Figura 13 *derecha*). Sin embargo, no se pudo asignar inequívocamente las señales de los residuos B y C, el solapamiento de los protones amídicos de estos residuos no permite distinguirlos (Figura 9). La asignación que se muestra en la figura 13 asume que el compartimiento de los átomos de flúor de estos residuos es similar al observado en el análogo de quitobiosa, en el que los protones más desapantallados pertenecen al extremo no-reductor.

	análogo fluorado de N,N'-diacetil quitobiosa		análogo fluorado de N,N',N''-triacetil quitotriosa	
	$^2J_{\text{FF}} \text{ (Hz)}$	$^2J_{\text{HF}} \text{ (Hz)}$	$^2J_{\text{FF}} \text{ (Hz)}$	$^2J_{\text{HF}} \text{ (Hz)}$
Residuo A	303	53,8	303	53,8
Residuo B	304	53,8	304	53,8
Residuo C	---	---	305	53,8

Tabla 1: Constantes de acoplamiento homonuclear ^{19}F - ^{19}F , $^2J_{\text{FF}}$, y constantes de acoplamiento heteronuclear ^1H - ^{19}F , $^2J_{\text{HF}}$, del grupo difluoroacetamido de los derivados de N,N'-diacetil quitobiosa y N,N',N''-triacetil quitotriosa con etiquetas difluoroacetamida.

4.2. Análogos fluorados de quitobiosa y quitotriosa en estado asociado

La interacción de estos dos ligandos con WGA se siguió mediante experimentos monodimensionales de ^1H y ^{19}F RMN, apoyados por experimentos STD y DOSY.

En el caso del disacárido, los experimentos mostraron su interacción con WGA. El proceso de reconocimiento se deduce de la observación de cambios en prácticamente todas las señales de ^1H RMN, ya que la mayoría de ellas se ensancha por la presencia de la proteína (Figura 14). La anchura de línea de las señales depende de numerosos factores, incluyendo la presencia de intercambio químico. En este caso, dado que, en el estado asociado, el disacárido (510 Da) está asociado a una entidad de alto peso molecular (WGA, 36 kDa), la relajación transversal se acelera y las señales se ensanchan, especialmente aquellas que corresponden a átomos próximos al receptor. Este análogo fluorado mostró mayor ensanchamiento en las señales del extremo no-reductor. Por ejemplo, protón CHF_2 del residuo B (no-reductor), se ensancha

significativamente, en contraste con el del residuo A, que sólo se perturba ligeramente (Figura 14, *izquierda*). Es decir, el reconocimiento molecular con la WGA se da principalmente a través del extremo no-reductor.

El mismo efecto se observó en los espectros de ^{19}F . Este efecto depende de la relación molar entre el ligando y la proteína (Figura 15), como es de esperar para un proceso de intercambio químico que es rápido-intermedio en la escala de tiempos de los desplazamientos químicos. Las señales de ^{19}F de B mostraron mayor ensanchamiento que las de A.

La interacción también se analizó mediante experimentos ^1H -STD. Los protones que recibieron mayor saturación fueron los del extremo no-reductor, aunque el residuo A también recibió saturación. De hecho, incluso la cadena alifática anclada en posición anomérica del residuo A recibe cierta saturación (Figura 16).

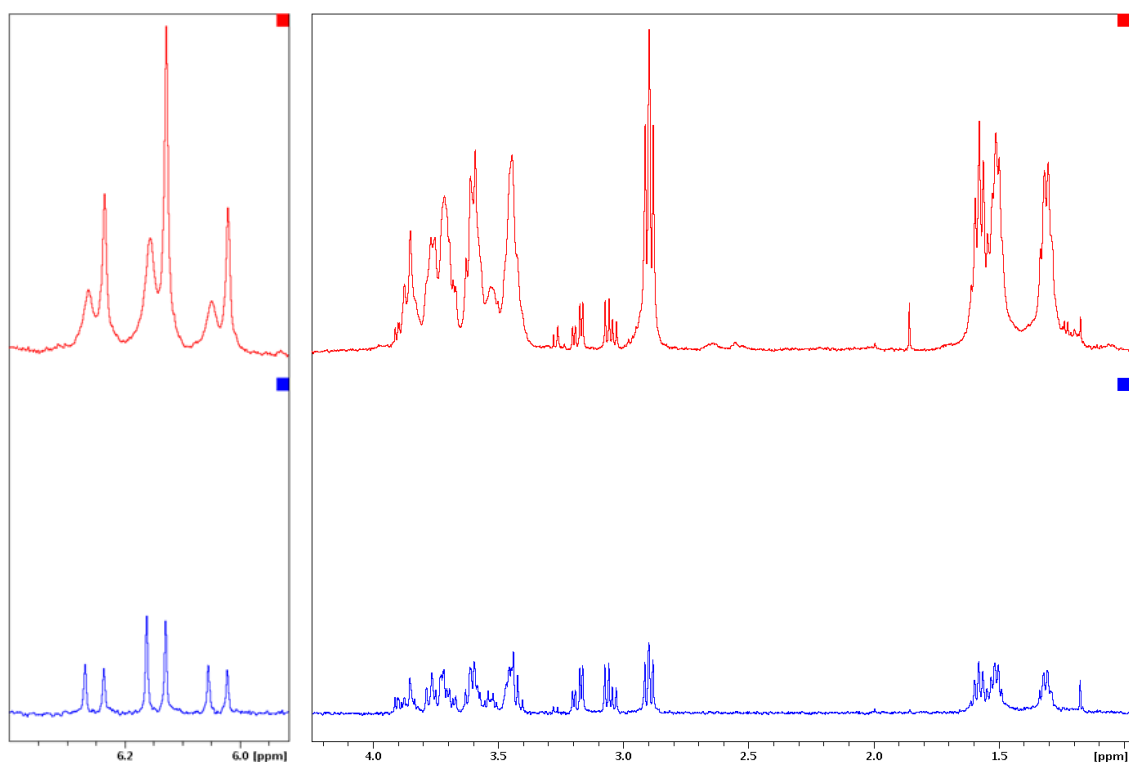


Figura 14: Interacción del derivado de N,N'-diacetil quitobiosa con etiquetas difluoroacetamida frente a WGA. *Espectro azul*, espectro ^1H del disacárido en ausencia de proteína. *Espectro rojo*, espectro ^1H en presencia de WGA. A la *izquierda* se muestra la región de los protones del grupo CHF_2 y, a la *derecha*, la región alifática. Estos espectros se hicieron en presencia de una referencia interna, la p-fluoro-L-fenilalanina. Este control negativo, que se observa entre 3-3,2 ppm; no sufre ensanchamiento en presencia de WGA.

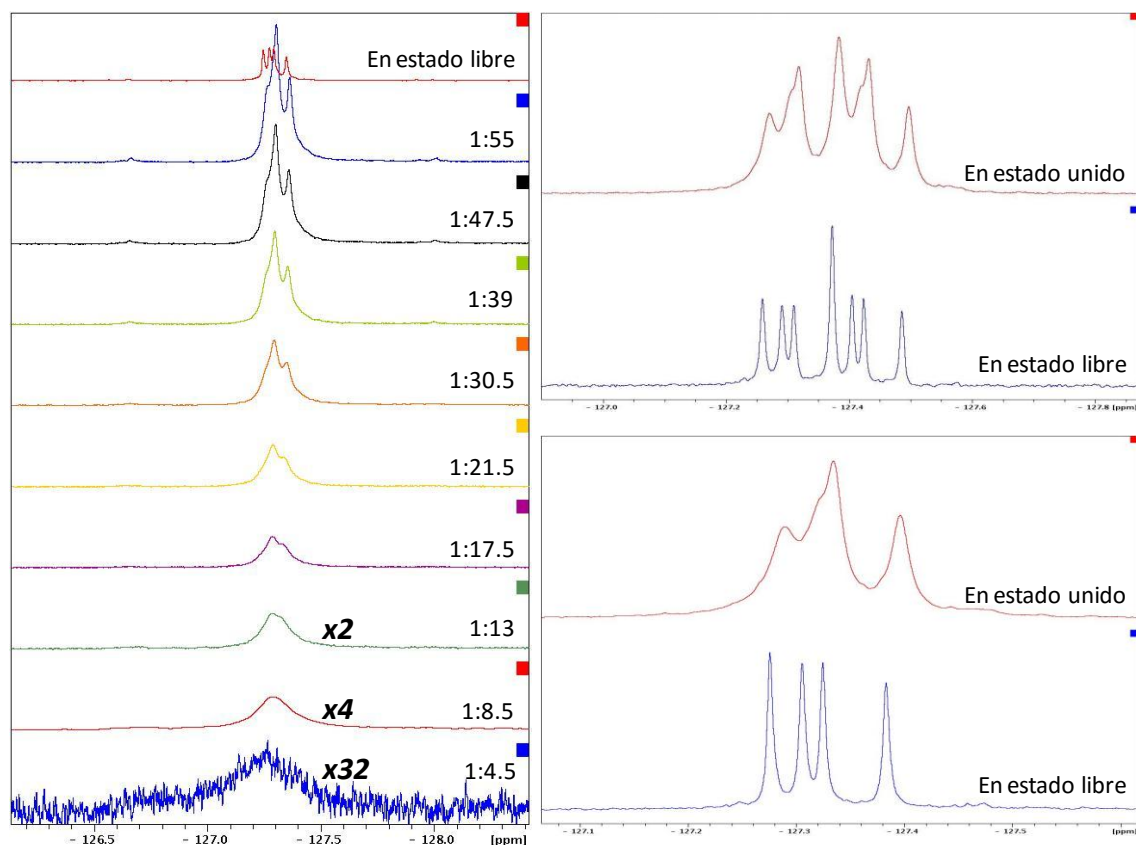


Figura 15: Interacción del derivado de N,N'-diacetil quitobiosa con etiquetas difluoroacetamida frente a WGA. *Izquierda*, espectros ^{19}F del disacárido en ausencia y en presencia de la proteína a distintas relaciones molares (desde 1:4.5 a 1:55). *Derecha*, comparación entre de las señales del estado unido y del estado libre en los espectros de ^{19}F sin desacoplamiento de protón (*arriba*) y con desacoplador de protón. (*abajo*).

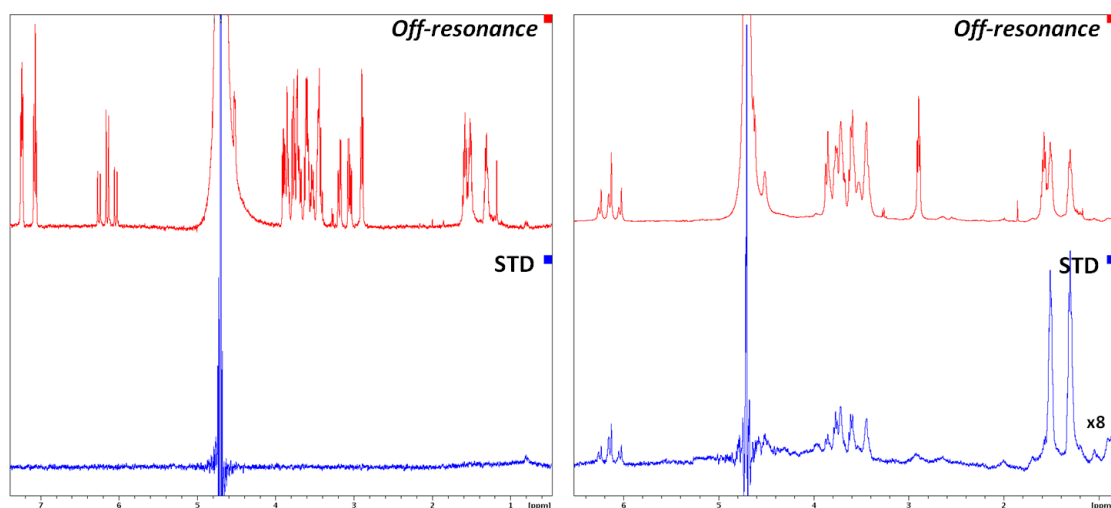


Figura 16: Experimentos ^1H -STD del derivado de N,N'-diacetil quitobiosa con etiquetas difluoroacetamida frente a WGA (tampón fosfato 50mM a pH 6.0, 298K). *Izquierda*, experimento control irradiando a 0.29 ppm en ausencia de la proteína. *Derecha*, experimento STD irradiando a 0.29 ppm en presencia de WGA.

Finalmente, también se realizaron experimentos de ^{19}F -STD. Se utilizaron dos distintas secuencias de pulsos, secuencia *FSTD* y secuencia *STDreF*.²⁵⁹ La secuencia *FSTD* utiliza la transferencia directa entre los ^1H de la proteína y los núcleos de ^{19}F del ligando. En cambio, la secuencia *STDreF* utiliza la saturación recibida por los protones del ligando para transferirla a los átomos de ^{19}F a través de las constantes de acoplamiento heteronucleares ^1H - ^{19}F . Los resultados de la secuencia *FSTD* no fueron satisfactorios. Sin embargo, los experimentos *STDreF* sí confirmaron la existencia de interacción preferente con el residuo B (Figura 17). El espectro *off-resonance* tendría que tener la misma apariencia que un espectro de flúor con desacoplamiento de ^1H durante la adquisición (Figura 17), pero se da una clara distorsión debida al fuerte acoplamiento entre los átomos de flúor de un mismo residuo de GlcNAc.

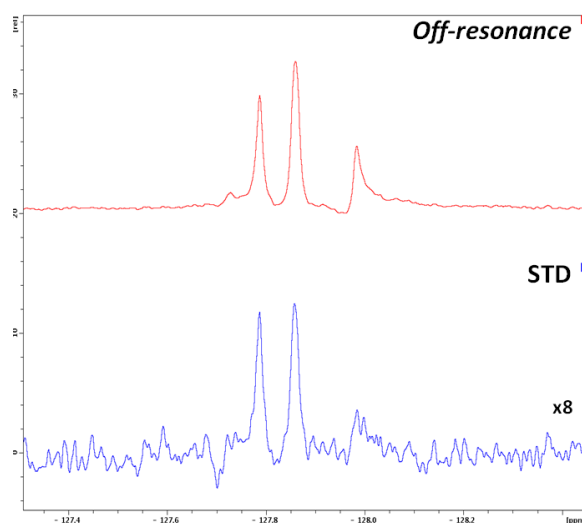


Figura 17: Experimentos ^{19}F -STD del derivado de N,N'-diacetil quitobiosa con etiquetas difluoroacetamida frente a WGA con la secuencia de pulsos *STDreF*²⁵⁹ (tampón fosfato 50mM a pH 6.0, 298K).

La interacción del trisacárido se analizó de manera análoga. El análisis de los espectros demostró que la interacción tenía lugar principalmente a través de sus anillos B y C (Figura 18). De hecho, el protón difluoroacetamido del residuo A no sufrió ensanchamiento apreciable a ninguna relación molar ligando:WGA (Figura 18 *izquierda*). Las señales de ^{19}F sufrieron un ensanchamiento más pronunciado que las de ^1H . Por ejemplo, los protones difluoroacetamido prácticamente no están afectados por la presencia de la lectina usando una relación molar de 1:100. Sin embargo, en estas condiciones experimentales, todas las señales de ^{19}F muestran una anchura de línea sensiblemente mayor que en el estado libre. De nuevo, las señales de ^{19}F del residuo A fueron las menos perturbadas por la presencia de la proteína (Figura 18 *derecha*).

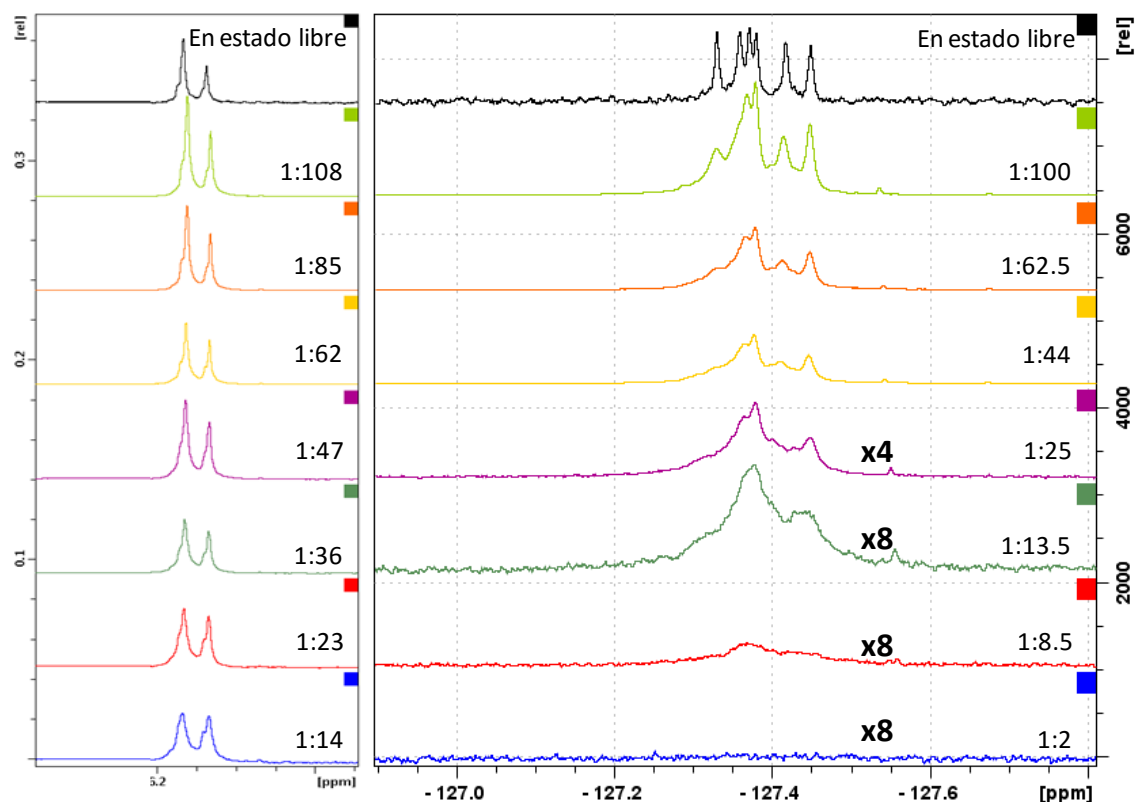


Figura 18: Interacción del derivado de N,N',N''-triacetil quitotriosa con etiquetas difluoroacetamida frente a WGA. *Izquierda*, región de los protones difluoroacetamida en los espectros ^1H desacoplando flúor durante la adquisición, usando distintas relaciones molares *proteína:ligando*. *Derecha*, espectros ^{19}F desacoplando protón durante la adquisición, usando distintas relaciones molares *proteína:ligando*.

También se realizaron espectros STD, tanto ^1H -STD como ^{19}F -STD, con el análogo fluorado de N,N',N''-triacetil quitotriosa, pero no se consiguió mostrar interacción de este ligando con estos experimentos como se había observado en el análogo fluorado de N,N',N''-triacetil quitobiosa. En cambio, la quitotriosa mostró mayor ensanchamiento que la quitobiosa en los espectros monodimensionales de protón y flúor. Esta observación experimental está de acuerdo con la existencia de una velocidad de intercambio más lenta en la escala de tiempos de relajación para este ligando. Esto indica que el análogo de quitotriosa se une con mayor afinidad que el de quitobiosa, como se ha observado previamente para los compuestos naturales^{57,260}.

Finalmente, también se realizaron espectros ^{19}F -DOSY (Diffusion Ordered Spectroscopy) que confirmaron la de interacción entre ambas entidades. Los espectros DOSY miden la velocidad de difusión translacional de la molécula observada, que será obviamente menor si se encuentra asociada a una macromolécula. El coeficiente del trisacárido es mayor en presencia de WGA (Figura 19). Si se comparan los coeficientes de difusión en estado libre y en una muestra cuya relación molar es 1:10, se observa un

incremento de 0.3 unidades logarítmicas debido a la interacción de alta afinidad de este ligando y la lectina.

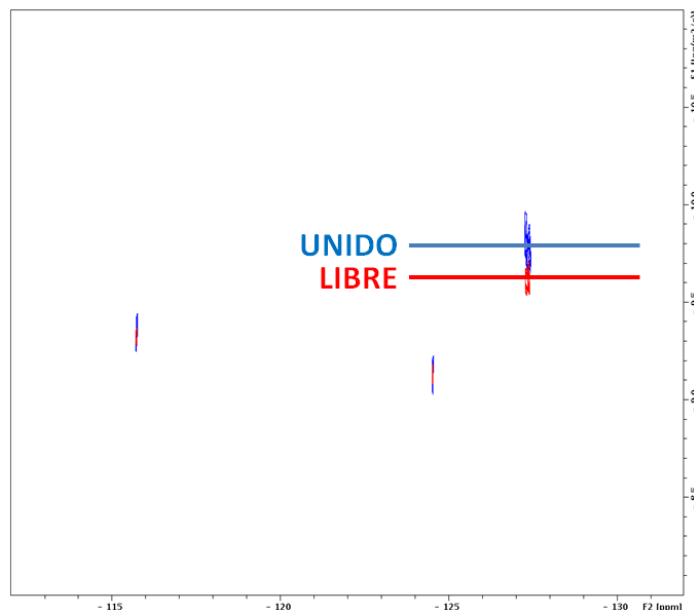


Figura 19: Experimentos DOSY del derivado de N,N',N''-triacetil quitotriosa con etiquetas difluoroacetamida. Superposición de dos espectros ^{19}F -DOSY de este análogo en ausencia (*espectro rojo*) y en presencia de WGA (*azul*). Estos espectros se hicieron en presencia de una referencia interna, la p-fluoro-L-fenilalanina. Este control negativo, cuya desplazamiento químico de flúor es -115,8 ppm, no cambia su coeficiente de difusión en presencia de WGA.

Todos los datos experimentales revelan que este motivo difluoroacetamida es una extraordinaria sonda dual para los ensayos de RMN, que puede ser utilizada en la investigación de interacciones proteína-carbohidrato de azúcares N-acetilados. Los experimentos de ^{19}F procuraran mayor sensibilidad en sistemas de baja afinidad, y los espectros de ^1H proveerán información estructural más detallada del epítipo.

4.3. Implicación del motivo difluoroacetamida en el proceso de reconocimiento

La interacción de estos azúcares que contienen el motivo difluoroacetamida ha sido demostrado por distintos experimentos de RMN, así que esta modificación química no impide la interacción de estas moléculas a los dominios heveína de WGA. Como el reconocimiento de moléculas que contienen residuos GlcNAc que tiene lugar en estos dominios ya ha sido investigado anteriormente²⁶¹⁻²⁶², se ha cuantificado el efecto de esta sustitución en las interacciones intermoleculares entre lectinas y ligandos con estos grupos.

La interacción se estudió mediante modelos simplificados que reproducen las interacciones de apilamiento (CH/π y CF/π) que tienen lugar en los complejos carbohidrato-lectina. Los cálculos *ab initio* se realizaron con el programa *Gaussian 09*. Se construyeron modelos optimizados de los complejos moleculares formados entre 4 moléculas modelo del grupo acetamido o sus análogos fluorados ($\text{CH}_3\text{-NH-CO-CF}_x\text{H}_{3-x}$) y sistemas modelo de residuos aromáticos (benceno y el p-metilfenol). Las distintas amidas se colocaron sobre los anillos aromáticos estableciendo interacciones CH/π o CF/π , para tener en cuenta ambas posibilidades de apilamiento (*stacking*).

El complejo $\text{CH}_3\text{-NH-CO-CF}_2\text{H}$ /benceno presenta las interacciones más estables cuando el *stacking* es de tipo CH/π . Es decir, con el enlace C-H apuntando hacia el plano aromático. En el complejo $\text{CH}_3\text{-NH-CO-CFH}_2$ /benceno también está favorecida la orientación de los enlaces C-H hacia el anillo. De hecho, la energía de este complejo es prácticamente la misma que en el grupo difluoroacetamida. En cambio, la orientación de los enlaces C-F de estas moléculas hacia el modelo aromático está desfavorecida, siendo la diferencia de energías superior a 2.5 kcal/mol entre la disposición con una interacción CH/π y la disposición con interacción CF/π (modelos N-Me-fluoroacetamido y N-Me-difluoroacetamido). Además, la interacción con el anillo aromático de estos grupos parcialmente fluorados es más favorable energéticamente de la que se da entre el anillo aromático y el metilo del grupo acetamido clásico. Por ejemplo, la interacción CH/π del grupo CF_2H es 1 kcal/mol más favorable que el *stacking* del grupo metilo, según los valores calculados para estos sistemas simplificados. Por otro lado, la interacción de *stacking* del grupo $-\text{CF}_3$ con el anillo aromático del benceno es menos favorable, ya este grupo sólo puede interaccionar a través de enlaces C-F que están penalizados energéticamente. La interacción CF/π del grupo $-\text{CF}_3$ es 1.6 kcal/mol menos estable que la interacción CH/π del grupo $-\text{CH}_3$ (Tabla 1). Los resultados de los cálculos para los complejos con p-metilfenol, como modelo de tirosina, fueron equiparables a los valores calculados para el modelo de $\text{CH}_3\text{-NH-CO-CF}_x\text{H}_{3-x}$ con benceno.

molécula acetamido ($\text{CH}_3\text{-NH-CO-CF}_x\text{H}_{3-x}$)	tipo de interacción <i>stacking</i>	E_{int} (Kcal/mol)	E_{BSSE} (Kcal/mol)	E_{corr} (Kcal/mol)
N-Me-acetamido ($\text{CH}_3\text{-NH-CO-CFH}_3$)	CH/π	-7,0	2,1	-4,8
N-Me-fluoroacetamido ($\text{CH}_3\text{-NH-CO-CFH}_2$)	CH/π	-8,9	3,3	-5,6
	CF/π	-6,2	3,3	-3,0
N-Me-difluoroacetamido ($\text{CH}_3\text{-NH-CO-CF}_2\text{H}$)	CH/π	-8,7	2,9	-5,7
	CF/π	-6,4	3,5	-2,9
N-Me-trifluoroacetamido ($\text{CH}_3\text{-NH-CO-CF}_3$)	CF/π	-6,7	3,6	-3,2

Tabla 2: Energías de interacción, calculadas mediante *Gaussian 09*, para los complejos entre el benceno y las 4 amidas modelo ($\text{CH}_3\text{-NH-CO-CF}_x\text{H}_{3-x}$). E_{int} , energía de interacción del complejo sin corrección; E_{BSSE} , valor de corrección de la energía de interacción del complejo; E_{corr} , energía de interacción aplicando la corrección; y E_{int} en H_2O , energía de interacción del complejo en solvente implícito.

La geometría de los distintos complejos corrobora el efecto de los valores energéticos de los distintos grupos acetamido en la posición relativa de los componentes de estos complejos. Las interacciones de apilamiento CH/ π tienen una geometría casi idéntica en todos los casos. Sólo se dan pequeñas diferencias entre los grupos que contienen algún átomo de flúor y el modelo que contiene el grupo -CH₃. En concordancia con los valores de energía, las distancias CH/ π para los grupos CF₂H y CFH₂ son 0.15 Å más cortas que estas distancias para el grupo -CH₃. Asimismo, las energías calculadas muestran que las interacciones CH/ π están favorecidas frente a las interacciones CF/ π y, consecuentemente, este hecho también se refleja en las distancias intermoleculares. Las distancias entre los átomos de hidrógeno y el centroide de los anillos aromáticos son 0.4-0.5 Å menores que las distancias de los átomos de flúor a los mismos.

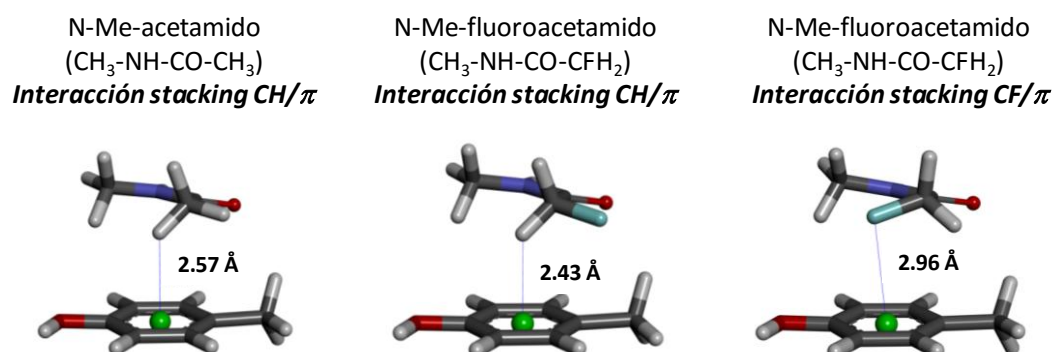


Figura 20: Geometrías de los complejos formados por p-metilfenol y dos acetamidas modelo, CH₃-NH-CO-CH₃ y CH₃-NH-CO-CFH₂. La distancia entre los núcleos implicados en la interacción de apilamiento y el centroide del anillo aromático se especifica en cada uno de los complejos.

4.4. Reconocimiento molecular de los análogos fluorados de quitobiosa y quitotriosa

Se han publicado diversas estructuras tridimensionales de rayos X de las isoformas de WGA en presencia de los distintos ligandos^{56,63,239,242}. El análisis de estos complejos también ha sido abordado mediante RMN. De hecho, trabajos anteriores ya han descrito la interacción de la quitobiosa y la quitotriosa frente a la heveína o un dominio de WGA^{245,263}. Estos estudios han permitido que se conozca detalladamente el proceso de reconocimiento de estos oligosacáridos en los dominios de heveína²⁶⁴.

En la interacción de la quitobiosa con un dominio heveína, el residuo GlcNAc del extremo no-reductor está localizado el sitio +1 del centro de reconocimiento. Este extremo no-reductor tiene una interacción CH/ π con el residuo aromático en posición

23, un enlace de hidrógeno con la Ser¹⁹, y un enlace de hidrógeno y apilamiento con la Tyr³⁰. El otro residuo de GlcNAc está localizado en el sitio +2 y su principal interacción es mediante una interacción CH/ π con el residuo en posición 21. Por otra parte, la quitotriosa muestra dos modos de unión diferentes con la heveína⁵⁸. El primer modo de unión, modo A, solapa con el de la quitobiosa en el centro de reconocimiento. El trisacárido coloca su extremo no-reductor en el sitio +1, mientras que los residuos en posición intermedia y en el extremo reductor ocupan los sitios +2 y +3, respectivamente. Por lo tanto, el residuo del extremo reductor en el modo A de unión presenta muy pocos contactos con la lectina. Sin embargo, en el modo de unión alternativo (B), el residuo intermedio de la quitotriosa se coloca en el sitio +1 mientras que el extremo reductor, en el sitio +2. De hecho, las interacciones del extremo no-reductor en el modo A son reproducidas por el residuo intermedio de GlcNAc en el modo B. Consecuentemente, el residuo GlcNAc del extremo no-reductor se coloca en el sitio -1 en este modo de unión.

Además de los cálculos *ab initio* mencionados en el apartado anterior, se realizaron cálculos de dinámica molecular de los análogos difluorados de la quitobiosa y la quitotriosa asociados a un dominio heveína. Para este análisis computacional se seleccionó la estructura con código PDB 1T0W como dominio modelo de heveína. El punto de partida de la dinámica molecular tuvo en cuenta las estructuras publicadas para los oligosacáridos naturales y los resultados de los cálculos de *Gaussian* mencionados anteriormente. Los cálculos se realizaron usando el campo de fuerzas AMBER²⁶⁵ con el software Maestro¹⁸².

El complejo del análogo fluorado de quitobiosa fue estable en el transcurso de la dinámica y mostró un modo de unión idéntico al publicado para la molécula no fluorada. Además, las interacciones CH/ π del grupo difluoroacetamido predominaron frente a las interacciones CF/ π de este grupo. El motivo difluoroacetamido del extremo no-reductor de este oligosacárido establece interacciones CH/ π con la Tyr³⁰, mientras que este mismo motivo, en el residuo reductor, está orientado hacia el disolvente (Figura 20). Los resultados con el análogo fluorado de quitotriosa mostraron que ambos modos de unión eran estables. En el modo de unión A, el grupo difluoroacetamido del extremo no-reductor establece una interacción CH/ π con la Tyr³⁰. En cambio, en el modo de unión B, es el grupo difluoroacetamido del GlcNAc en posición intermedia (Figura 21) el que se coloca en la posición adecuada. En resumen, los resultados de las dinámicas moleculares corroboran que la sustitución CHF₂ no afecta a procesos de reconocimiento de estos oligosacáridos. Estos resultados computacionales están de acuerdo con las observaciones experimentales de RMN descritos anteriormente, que mostraron que las señales de ¹H y ¹⁹F de los residuos del extremo no-reductor, y del residuo central del análogo de quitotriosa, se afectan más fuertemente que las señales del extremo reductor.

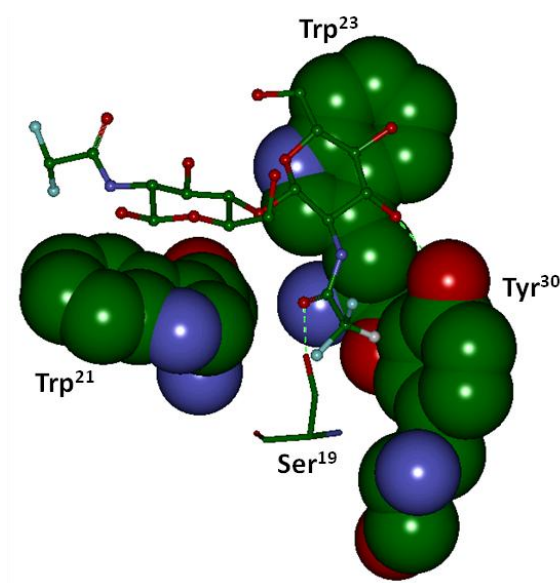


Figura 21: Representación esquemática del análogo fluorado de quitobiosa en el sitio de reconocimiento de la heveína.

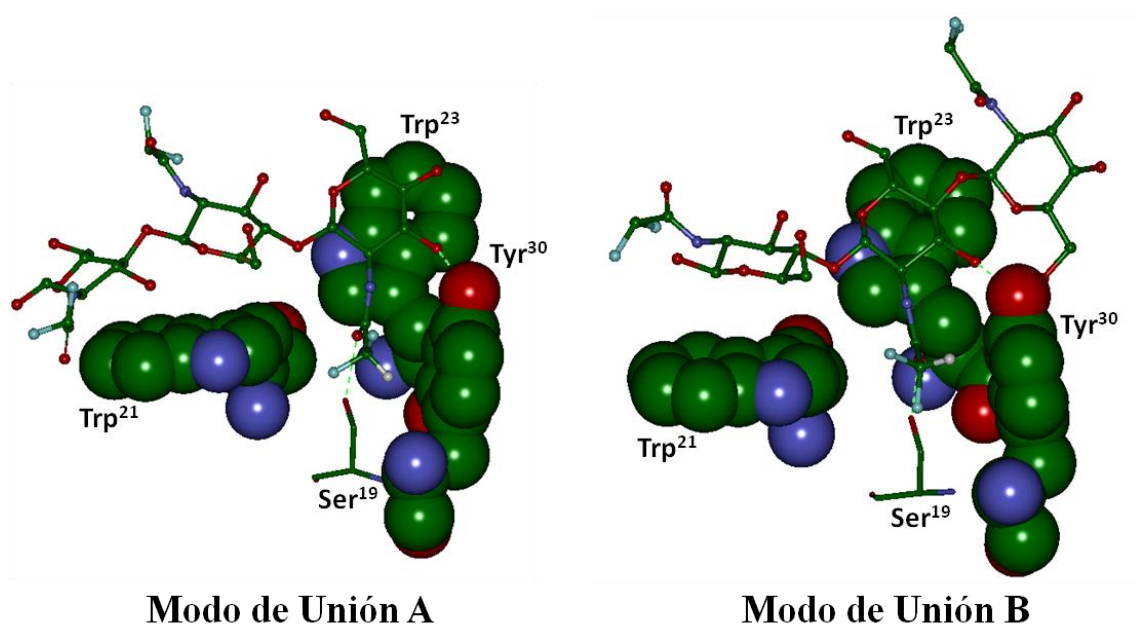


Figura 22: Representación esquemática del análogo fluorado de quitotriosa en el sitio de reconocimiento de la heveína. El ligando conserva ambos modos de unión durante la simulación de dinámica molecular. En el *modo de unión A*, la Tyr³⁰ presenta CH/ π con el grupo difluoroacetamido del extremo no-reductor. En el *modo B*, la interacción tiene lugar con el grupo difluoroacetamido del residuo en posición intermedia.

4.5. Conclusiones: Posibilidades de uso de esta sonda difluoroacetamida para detectar interacciones

Este estudio ha demostrado la utilidad del grupo difluoroacetamida para su uso como sonda en el análisis de interacciones carbohidrato-proteína. Aunque este grupo ya ha sido empleado en química médica²⁶⁶, la utilización del motivo difluoroacetamida como sensor en procesos de reconocimiento molecular de carbohidratos no había sido demostrado todavía. Esta sonda permite la detección fácil y la caracterización específica de procesos de reconocimiento mediante RMN. Puesto que esta etiqueta contiene átomos de ^1H y ^{19}F , la interacción de los compuestos que la contengan podrá ser analizada mediante experimentos de ^1H o de ^{19}F RMN. Los cálculos *ab initio* muestran que los dos flúores geminales mejoran la interacción CH/π , ya que polarizan el enlace C-H contiguo, y determinan que exista una orientación favorecida energéticamente. Estas observaciones revelan la utilidad de esta estrategia para el diseño de ligandos en procesos de reconocimiento molecular en los que estén implicados residuos GlcNAc, GalNAc o NeuNAc.

4.6. Materiales y métodos experimentales

Los compuestos análogos de N,N'-diacetil quitobiosa y N,N',N''-triacetil quitotriosa con las etiquetas difluoroacetamida fueron sintetizados por el grupo del Dr. Niels-Christian Reichardt de CIC BiomaGUNE, San Sebastián.

La aglutinina de germen de trigo (*Wheat Germ Agglutinin*, WGA) fue adquirida a Sigma-Aldrich. La preparación comercial se filtró con PES VIVASPIN 6 con un paso límite de 10 kDa y se disolvió en un tampón fosfato 50mM a pH 6.0 en D_2O para realizar los ensayos con los ligandos.

4.6.1. Resonancia Magnética Nuclear

Los experimentos de RMN se adquirieron a 298 K en un espectrómetro Bruker AV500 utilizando una sonda SEF ("Fluor selective"), con una bobina interna optimizada para la observación de ^{19}F . El proceso de asignación de los espectros de ^1H se efectuó usando muestras de concentración 2mM, en D_2O , en tampón fosfato 50mM, a pH 6.0 en D_2O , y en el mismo tampón usando $\text{H}_2\text{O}:\text{D}_2\text{O}$ (90:10) para detectar los protones HN. Se utilizaron secuencias estándar de experimentos TOCSY (*dipsigpph19*, con dos tiempos de mezcla 20ms y 60ms), experimentos NOESY (*noesygpph19*, con dos tiempos de mezcla 300ms y 500ms), experimentos NOE selectivos (*selnogp*, tiempo de mezcla 500ms) y experimentos HSQC (*hsqcetgp*). Los espectros de ^{19}F se asignaron mediante experimentos de correlación escalar homonuclear ^{19}F - ^{19}F (modificación de una secuencia estándar de Bruker) y heteronuclear ^{19}F - ^1H (*hfcofqqn*).

Los experimentos en presencia de la proteína se realizaron en tampón fosfato 50mM a pH 6.0 en D_2O , con concentraciones de ligandos y la lectina dependientes de los

requerimientos experimentales. Los experimentos ^1H (*p3919p*) y ^{19}F (*zgflqn* y *zgfhigqn*) fueron obtenidos con secuencias estándar de Bruker. Los experimentos ^1H STD (*std2*) se realizaron con un tiempo de saturación de 2 segundos y frecuencia off-resonance a 100 ppm, mientras que la frecuencia on-resonance se fijó a 0.29 ppm. Los experimentos ^{19}F STD se ejecutaron mediante las secuencias publicadas por los grupos de investigación del CIB y CICbioGUNE²⁵⁹, utilizando parámetros análogos a los ^1H STD. Los espectros ^{19}F DOSY se adquirieron con una secuencia de difusión modificada, a la que se añadió un modulo para el desacoplamiento heteronuclear con los protones (composite pulse decoupling CPD) durante la adquisición

4.6.2. Cálculos *ab initio* mediante el programa Gaussian 09

Los cálculos *ab initio* fueron realizados por D. Antonio Franconetti en el departamento de *Química Orgánica* de la Universidad de Sevilla, siguiendo el protocolo descrito anteriormente. Los cálculos fueron realizados con el programa *Gaussian 09*²⁶⁷, y ninguno de los complejos fue sometido a restricciones. Las optimizaciones geométricas se hicieron mediante el método funcional meta híbrido M06-2X (*hybrid meta exchange-correlation functional M06-2X*)²⁶⁸⁻²⁷⁰ a un nivel de teoría 6-31G(d,p). El análisis de frecuencia vibracional se realizó usando el mismo nivel de teoría para asegurar que la geometría obtenida era un mínimo energético. Además, el efecto del solvente, agua, se estimó aplicando un modelo continuo mediante la variante IEFPCM (*integral equation formalism polarizable continuum model*).

La energía de las interacciones CH/ π y CF/ π se calcularon usando las fórmulas que se dan más adelante. La energía neta de estas interacciones se obtuvo sustrayendo la energía calculada para cada monómero de la energía del complejo.

$$\Delta E_{\text{int.stacking}} = E_{\text{complejo}} - E_{\text{acetamida}} - E_{\text{aromático}}$$

Estas energías se corrigieron considerando el error de superposición de los conjuntos de bases (BSSE)²⁷¹. Finalmente, se realizaron cálculos rigurosos (con correcciones *counterpoise*), sin optimización, para cada una de las geometrías específicas.

$$\Delta E_{\text{corregida}} = \Delta E_{\text{int.stacking}} + E_{\text{BSSE}}$$

4.6.3. Dinámica molecular

Existen distintas estructuras de WGA y de dominios heveína descritas mediante difracción de rayos X o de RMN (códigos *PDB*: 1K7U²⁴⁴, 7WGA⁵⁶, 2UVO²³⁹, 1Q9B²⁷², 1T0W²⁷³). En este estudio se ha seleccionado un domino modelo de heveína para realizar los cálculos de dinámica molecular (1T0W). Esta estructura inicial se modificó reemplazando el grupo acetamida de los ligandos por el grupo difluoroacetamida. La orientación se escogió para conseguir una interacción estable CH/ π entre el hidrógeno de este grupo y la Tyr³⁰, de acuerdo con los cálculos *ab initio*. Los cálculos se realizaron con el software Maestro,¹⁸² usando el campo de fuerzas AMBER²⁶⁵.

5. Inhibidores de β -galactosidasa a través de modificaciones en anillo aglicónico

β -galactosidasa de *Escherichia coli* es una proteína que ha sido empleada como modelo de enzimas glicosidasas que siguen un mecanismo de doble desplazamiento con retención de la configuración en la posición anomérica del producto. Existen diversas estructuras de rayos X con diferentes ligandos en el *Protein Data Bank*. El análisis de las estructuras tridimensionales de este enzima, junto con diversos ensayos bioquímicos, ha permitido conocer el mecanismo de acción hidrolítico de esta glicosidasa. Las conclusiones se basaron adicionalmente en estudios funcionales con distintos mutantes en ciertos residuos y se corroboraron mediante el estudio mediante RMN de la conformación de glicomiméticos e inhibidores en su centro enzimático^{99,274}.

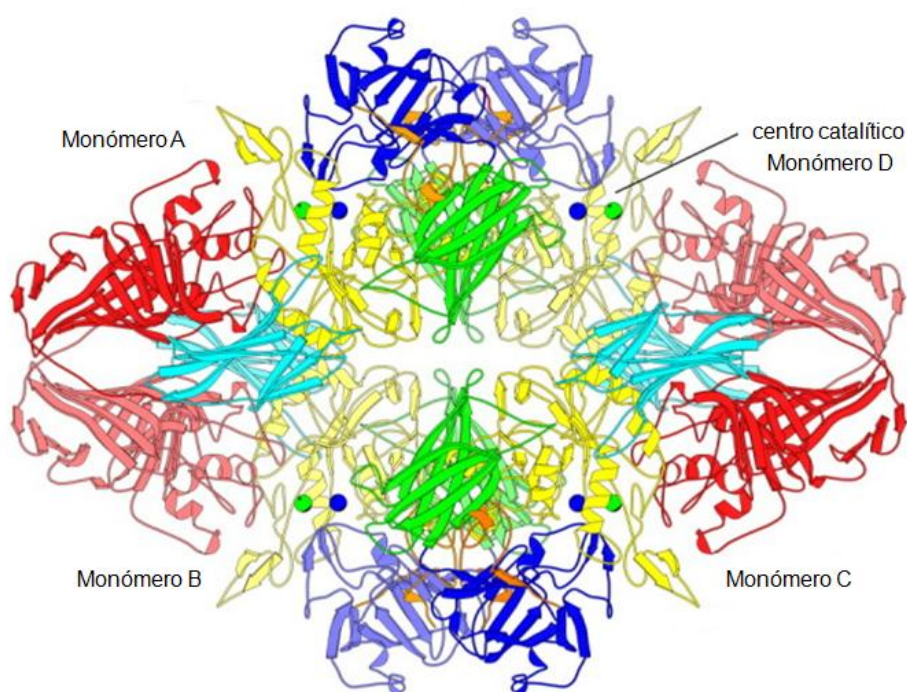


Figura 1: Estructura del homotetrámero de la β -galactosidasa de *E.coli*: Los cuatro monómeros se nombran de la A a la D, y los cinco dominios se diferencian por el color de sus estructuras secundarias (el dominio 1, azul; el dominio 2, verde; el dominio 3, amarillo; el dominio 4, cian; y el dominio 5, rojo). Además, los dos metales del dominio catalítico Na^+ y Mg^{2+} , se muestran como esferas de color verde y azul, respectivamente²⁷⁵.

El centro catalítico se encuentra en la cavidad de una estructura de *barril TIM*. El término *barril TIM* o *TIM-barrel* proviene del ejemplo, particularmente simétrico, de la enzima *Triose phosphate IsoMerase* (TIM),²⁷⁶ que comparte una estructura de hélices α y láminas β con la β -galactosidasa de *E. coli*. La β -galactosidasa (*lacZ*) de esta bacteria es una proteína homotetrámera de 464kDa (116 kDa por monómero), cuyas

subunidades constituyen 5 dominios. El tercer dominio conforma el *barril* α/β que contiene los residuos del sitio activo y mantiene homología de secuencia y estructural con toda una serie de enzimas glicosidasas de diversos orígenes y especificidades que forman la familia 2 de glicosidasas según la clasificación recogida en la base de datos CAZY (<http://www.cazy.org/>)²⁷⁷.

Mediante mutagénesis, se ha identificado que el residuo catalítico que produce el ataque nucleofílico sobre el ligando es el Glu⁵³⁷. Los otros residuos esenciales para la acción enzimática son el Glu⁴⁶¹, que interacciona con el ion magnesio del dominio 3 y que actúa con catalizador ácido en la etapa de formación del intermedio *glicosil enzima* y como catalizador básico en la segunda etapa de hidrólisis del intermedio, y la Tyr⁵⁰³ que ha sido sugerida como el residuo ácido en segunda etapa²⁷⁸⁻²⁸⁰. Además de estos residuos catalíticos, existen otros esenciales para el reconocimiento molecular, entre los que cabe destacar el Trp⁵⁶⁸ y el Trp⁹⁹⁹. Estos dos residuos aromáticos establecen interacciones CH- π (*CH- π stacking interaction*) entre los residuos glicosídicos del ligando y los aminoácidos de la proteína. Concretamente, el Trp⁵⁶⁸ asume este tipo de interacciones entre los protones en posición 3, 4 y 5 del anillo de la piranosa del extremo no reductor, la galactosa, y el Trp⁹⁹⁹ con aquellos protones del aglicón, del otro lado del enlace glicosídico²⁸¹.

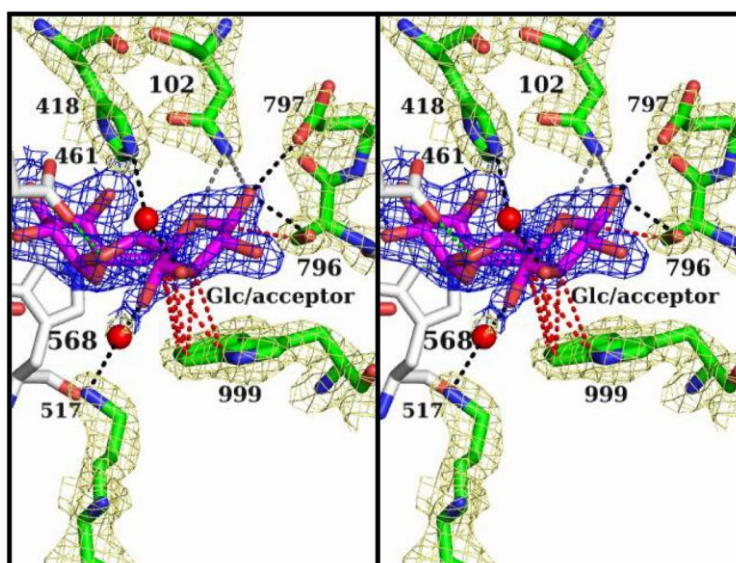
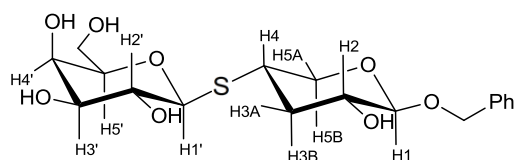


Figura 2: Estereograma del centro activo de la β -galactosidasa de *E.coli* mutante G794A con alolactosa (PDB: 4DUW). Los números referencian algunos de los aminoácidos importantes en el centro activo de la hidrolasa como Glu⁴⁶¹, Trp⁵⁶⁸ o Trp⁹⁹⁹, y otros residuos con interacciones presentes en esta estructura tridimensional. Líneas discontinuas rojas son interacciones CH- π de la glucosa con el Trp⁹⁹⁹, y las negras y las grises son los puentes de hidrógeno con otros residuos del bolsillo de la enzima.²⁸²

La proteína genera distorsiones en la conformación del ligando que pueden ser visualizadas por RMN. El análisis conformacional de O-glicósidos no puede ser realizado con la proteína nativa debido a que los hidroliza. Sin embargo, el uso mutantes en su residuo nucleófilo, Glu⁵³⁷ (mutante *E537Q*), permitió identificar la distorsión de la silla ⁴C₁ del residuo de galactosa en el extremo no reductor. Esta conformación semi-silla de alta energía, que la galactosa no adopta en disolución, es producida por el enzima y se asemeja al estado de transición propuesto para la hidrólisis, el ion oxocarbenio²⁸³. Otros experimentos de RMN con la proteína nativa se realizaron con análogos de sustrato, como C-glicósidos o S-glicósidos, ya que estos no sufren el ataque nucleófilo. Por el contrario, estos ligandos, miméticos de lactosa, en los que el efecto anomérico es inexistente o está muy minimizado, son conformacionalmente menos rígidos que los O-glicósidos y se pueden adaptar al sitio activo del enzima modificando el ángulo de torsión ϕ , alrededor del enlace glicosídico hacia una conformación *anti* menos poblada en solución en vez de forzar la distorsión de la silla del extremo no reductor. Es decir, las deformaciones del sustrato y sus análogos son interpretables en el sentido de que esta hidrolasa selecciona conformaciones de alta energía como primer paso del proceso de reconocimiento molecular¹¹⁶⁻¹¹⁷ para avanzar en el camino hacia el estado de transición de la reacción.

En este capítulo se presenta el análisis del reconocimiento molecular de una familia de nuevos inhibidores con estructura de S-galactósido y aglicón de tipo pentosa, preparados por el grupo del Prof. O. Varela en Buenos Aires: el bencil 3-deoxy-4-*S*-(β -D-galactopyranosyl)-4-thio- β -D-*erythro*-pentopyranoside (inhibidor **1**) y el bencil 3-deoxy-4-*S*-(β -D-galactopyranosyl)-4-thio- β -D-*threo*-pentopyranoside (inhibidor **2**).²⁸⁴ Estos S-glicósidos muestran flexibilidad conformacional adicional en su extremo reductor y podrían adoptar distintas geometrías para adaptarse al sitio de reconocimiento de la enzima sin penalización energética importante (Figura 3).

Inhibidor 1



Inhibidor 2

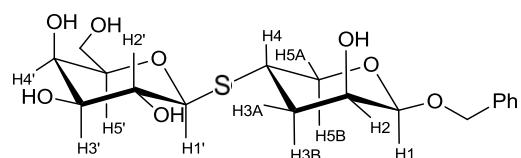


Figura 3: Estructuras de los dos análogos de sustrato de β -galactosidasa de *E.coli* estudiados en este capítulo, bencil 3-deoxy-4-*S*-(β -D-galactopyranosyl)-4-thio- β -D-*erythro*-pentopyranoside (inhibidor **1**, *izda.*) y el bencil 3-deoxy-4-*S*-(β -D-galactopyranosyl)-4-thio- β -D-*threo*-pentopyranoside (inhibidor **2**, *dcha.*).

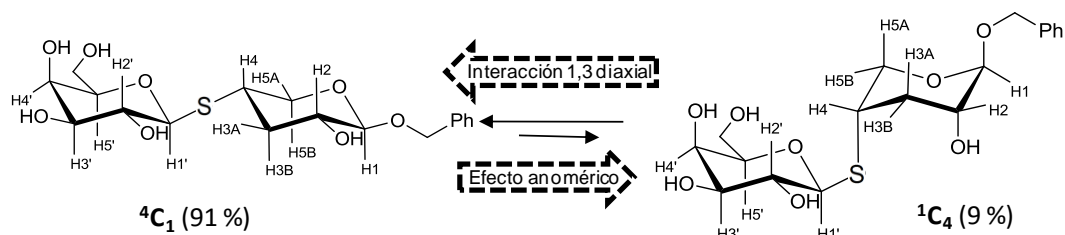
Los ensayos enzimáticos de estos inhibidores frente la β -galactosidasa de *E. coli* mostraron diferencias significativas, aunque estructuralmente sólo se diferencian en la estereoquímica de la posición C2 del anillo de pentopiranos. El inhibidor **1** muestra una moderada inhibición no-competitiva, cuya constante de inhibición es 800 μ M, mientras que el inhibidor **2** tiene un comportamiento competitivo, y más potente, con una $K_I=32\mu$ M.²⁸⁴ El análisis conformacional de estas moléculas en su estado libre y en su estado unido a esta glicosidasa, junto a la comparación de sus epítopos de reconocimiento, permitirán descifrar qué determina sus diferencias en inhibición.

5.1. Conformación de los tiogalactósidos **1** y **2** en estado libre

La conformación de ambos anillos (galactopiranos y pentopiranos) para ambos compuestos se determinó mediante un análisis clásico de constantes de acoplamiento (Tabla 1). Los valores para los anillos de galactosa de ambas moléculas son los esperados para una conformación típica de silla 4C_1 según la ecuación generalizada de Karplus propuesta por Altona, implementada en el programa MSpin (MestreLab Research, Santiago de Compostela, España)²⁸⁵. Sin embargo, para el anillo de pentopiranos, los valores experimentales no se pueden explicar con una única conformación y es necesario considerar un equilibrio conformacional entre las sillas 4C_1 y 4C_1 (Figura 4).

El anillo de 3-desoxi-*erythro*-pentopiranos de **1** adopta principalmente una conformación 4C_1 (>90%) con una pequeña contribución de la silla 1C_4 (<10%). Las poblaciones calculadas para ambas sillas nos permiten estimar que su diferencia energética es de 1.4 kcal/mol. Además, la existencia de una elevada preferencia conformacional por la silla 4C_1 es corroborada por la observación de una constante a larga distancia, $^4J_{H3A,H5A} = 2$ Hz. Esta constante sugiere una geometría tipo W para estos dos protones en disposición ecuatorial, que sólo es posible en la conformación 4C_1 . En cambio, **2** muestra una distribución igualada entre ambas sillas (Figura 6).

Inhibidor 1



Inhibidor 2

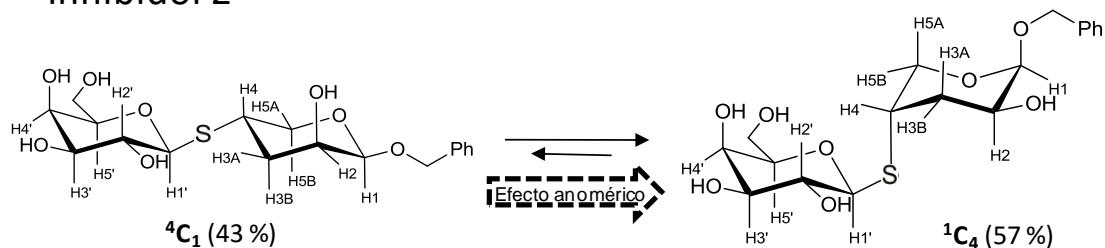
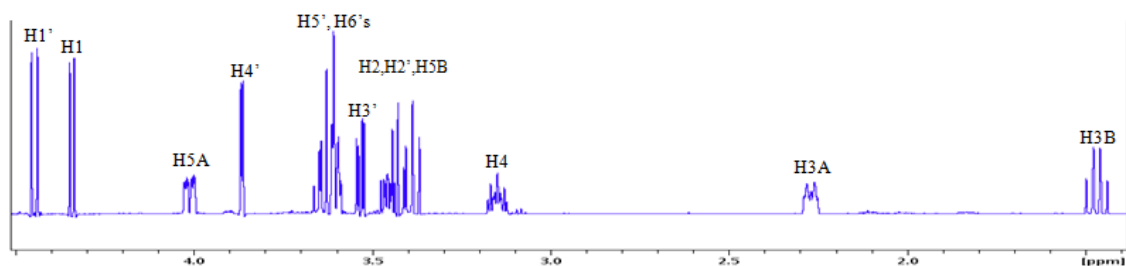


Figura 4: Equilibrio conformacional del anillo de pentopiranososa en los inhibidores 1 y 2 en disolución. Entre paréntesis se muestran las poblaciones calculadas.

$(^3J_{HH})$	1 1C_4 silla ($^3J_{HH}$ calculadas)	1 4C_1 silla ($^3J_{HH}$ calculadas)	1 $^3J_{HH}$ Experimentales	2 1C_4 silla ($^3J_{HH}$ calculadas)	2 4C_1 silla ($^3J_{HH}$ calculadas)	2 $^3J_{HH}$ Experimentales
H1-H2	2.0	6.5	7.4	3.1	0.9	2.6
H2*-H3A	3.3	5.1	4.7	10.9	3.4	7.8
H2*-H3B	2.7	10.7	10.7	4.8	2.9	4.1
H3A-H4	5.1	3.6	4.4	4.6	3.6	4.0
H3B-H4	1.9	11.9	10.8	2.1	11.9	8.7
H4-H5A	2.7	4.4	4.4	2.8	4.4	3.2
H4-H5B	1.0	11.3	11.4	1.0	11.3	5.7

Tabla 1: Constantes de acoplamiento vecinales, experimentales y calculadas, para los anillos de pentopiranososa de ambos inhibidores (Hz). La stereoquímica en C2 es diferente entre **1** y **2** (*).

INHIBIDOR 1



INHIBIDOR 2

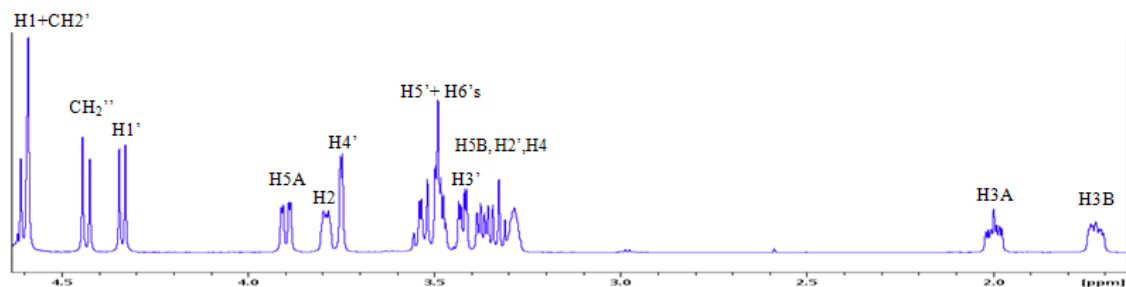


Figura 5: Asignación de los espectros monodimensionales de ambos inhibidores. En el espectro del inhibidor **1** no se muestran ni los protones del grupo CH₂ del bencilo (4.75 y 4.6 ppm) ni los aromáticos (multiplete a 7.35 ppm). En el espectro del inhibidor **2** se han señalado los dos protones del grupo CH₂, CH₂' y CH₂'', pero no se muestra el multiplete de los protones aromáticos a 7.35 ppm.

El efecto anomérico actúa en ambos inhibidores favoreciendo energéticamente la conformación ¹C₄ de la silla del residuo del extremo reductor, colocando su sustituyente de la posición anomérica en disposición axial. Sin embargo, la pentopiranososa tipo 3-desoxi-*erythro*-pentopiranososa de **1** adopta mayormente una conformación ⁴C₁, debido a que en la conformación ¹C₄ se genera una interacción 1,3-diaxial desestabilizante entre el hidroxilo en posición C2 y el azufre del enlace entre los anillos de piranosas. Esta disposición 1,3-diaxial no ocurre en la conformación ¹C₄ en el caso de la 3-desoxi-*threo*-pentopyranoside de **2**. El valor energético de esta interacción de repulsión puede ser obtenido a partir de la distribución de poblaciones de esta pareja de inhibidores. La diferencia energética del equilibrio entre la conformación silla ⁴C₁ y la silla ¹C₄ de **1** es 1.4 kcal/mol, mientras que **2** presenta un valor de -0.1 kcal/mol. Consecuentemente, esta interacción 1,3-diaxial (ΔΔG) tiene un valor estimado del orden de 1.5 kcal/mol (Figura 4).

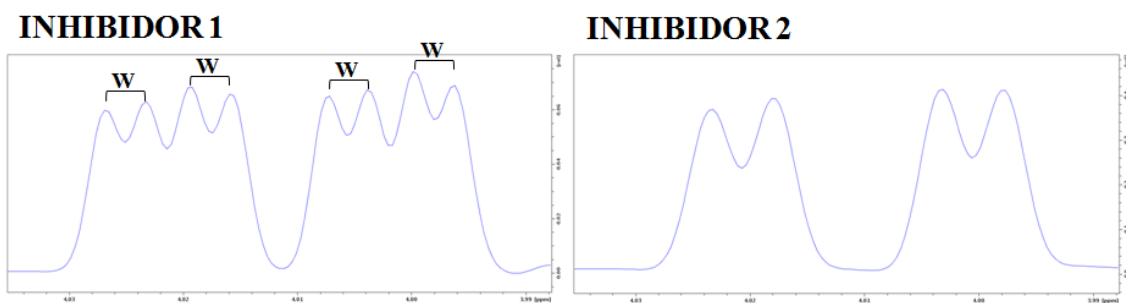


Figura 6: Comparación del patrón de acoplamiento del protón H5A de ambos inhibidores. El protón H5A del inhibidor **1** muestra claramente la constante a largo alcance tipo W, mientras que el mismo protón en el inhibidor **2** sólo muestra las constantes vecinales a tres enlaces.

En un segundo paso, se caracterizó la conformación de la silla de pentopiranososa y la conformación alrededor del enlace glicosídico, ϕ (H1'-C1'-S-C4) y ψ (C1'-S-C4-H4) mediante experimentos NOESY bidimensionales. El análisis de los NOEs intrarresiduales del anillo de pentopiranososa confirmaron las conclusiones del análisis de sus constantes de acoplamiento. En el caso del inhibidor **1** se observaron los NOEs característicos de la silla 4C_1 , H1-H3B y H3B-H5B, sin embargo no se observaron los picos NOE intrarresiduales esperados para la conformación 1C_4 , coincidiendo con el bajo peso poblacional calculado mediante el análisis de las constantes de acoplamiento. Los NOE intrarresiduales del anillo de pentopiranososa en el inhibidor **2** también están de acuerdo con el análisis previo de las constantes de acoplamiento, que indicaba la existencia de las dos conformaciones en el anillo de pentopiranososa para justificar los valores de las constantes de acoplamiento. Se observan los picos de cruce NOE entre los protones H3A-H5A y entre H3B-H5B, que sustentan respectivamente la silla 1C_4 y la 4C_1 .

El análisis de la conformación alrededor del enlace glicosídico es realizado únicamente mediante experimentos NOESY. Al no existir acoplamiento protón-protón interanillos, para el análisis conformacional se debe recurrir a una medida de efecto NOE entre protones de los distintos anillos complementado por una búsqueda por mecánica molecular de las geometrías conformacionales de mínima energía. Puesto que el residuo del extremo reductor puede adoptar dos conformaciones de silla diferentes, el análisis de Mecánica Molecular se realizó por duplicado. Los cálculos con el anillo de pentopiranososa en conformación de silla 4C_1 mostraron tres posibles mínimos energéticos ($\text{syn}\phi_{4C1}/\text{syn}\psi_{4C1}$, $\text{anti}\phi_{4C1}/\text{syn}\psi_{4C1}$ y $\text{syn}\phi_{4C1}/\text{anti}\psi_{4C1}$), mientras que la misma aproximación con este residuo en conformación 1C_4 llevó a dos posibles conformaciones ($\text{syn}\phi_{1C4}/\text{syn}\psi_{1C4}$ y $\text{anti}\phi_{1C4}/\text{syn}\psi_{1C4}$). En la figura 7 sólo se presentan los mapas de energía del inhibidor más potente, inhibidor **2**, ya que los dos inhibidores muestran el mismo comportamiento en los cálculos de Mecánica Molecular.

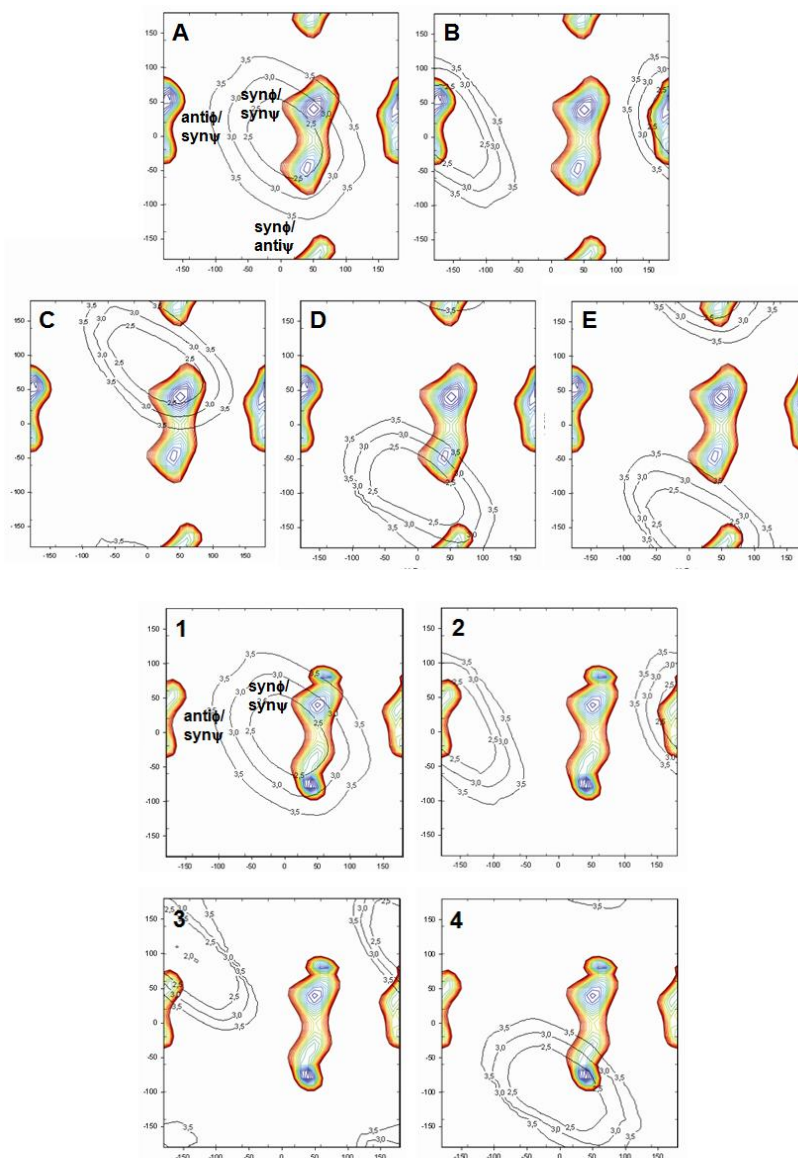


Figura 7: Representación de las distancias entre protones en curvas de nivel sobre los mapas energéticos de los cálculos de Mecánica Molecular del inhibidor **2** (eje X, ángulo diedro ϕ ; eje Y, ángulo diedro ψ). A-E corresponden a los mapas de energía con el anillo de pentopiranososa en conformación silla 4C_1 , 1-4 corresponden a los mapas con el anillo de pentopiranososa en conformación silla 1C_4 . Se representan las distancias interprotónicas susceptibles de generar NOEs ($< 3,5 \text{ \AA}$). A, H1'-H4; B, H2'-H4; C, H1'-H5A; D, H1'-H3A; E, H1'-H3B; 1, H1'-H4; 2, H2'-H4; 3, H1'-H5B; 4, H1'-H3A.

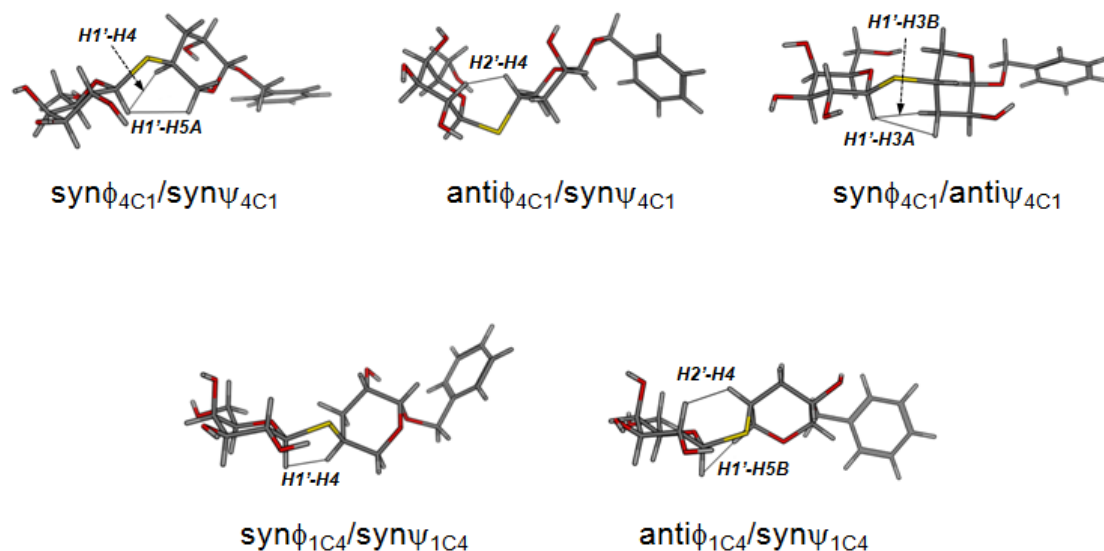


Figura 8: Representación tridimensional de los cinco mínimos energéticos del inhibidor **2**. *Arriba*, se representan los tres mínimos energéticos con el anillo de pentopiranososa en conformación silla 4C_1 ($\text{syn}\phi_{4C1}/\text{syn}\psi_{4C1}$, $\text{anti}\phi_{4C1}/\text{syn}\psi_{4C1}$ y $\text{syn}\phi_{4C1}/\text{anti}\psi_{4C1}$); *abajo*, se representan los dos mínimos energéticos con el anillo de pentopiranososa en conformación silla 1C_4 ($\text{syn}\phi_{1C4}/\text{syn}\psi_{1C4}$ y $\text{anti}\phi_{1C4}/\text{syn}\psi_{1C4}$). Se han señalado las distancias que describen cada una de las conformaciones.

En los espectros NOESY de **1** (Figura 9) el cófórmero $\text{syn}\phi_{4C1}/\text{syn}\psi_{4C1}$ es confirmado por la presencia de dos NOEs interresiduales para esta geometría, H1'-H4 y H1'-H5A (Figura 7, mapa A y C respectivamente). Además, la presencia de otro mínimo energético, el $\text{syn}\phi_{4C1}/\text{anti}\psi_{4C1}$, se demostró por la presencia de su contacto exclusivo, H1'-H3B (Figura 7, mapa E). Por último, la conformación $\text{anti}\phi_{4C1}/\text{syn}\psi_{4C1}$ no pudo ser identificada experimentalmente por el solapamiento existente entre su distintivo H2'-H4 interglucosídico (Figura 7, mapa B) y el esperado H2-H4 intrarresidual. Según este análisis, **1** puede adoptar en disolución hasta tres conformaciones alrededor de los enlaces glicosídico-aglicónico, donde el anillo de pentopiranososa adopta principalmente una conformación de silla 4C_1 .

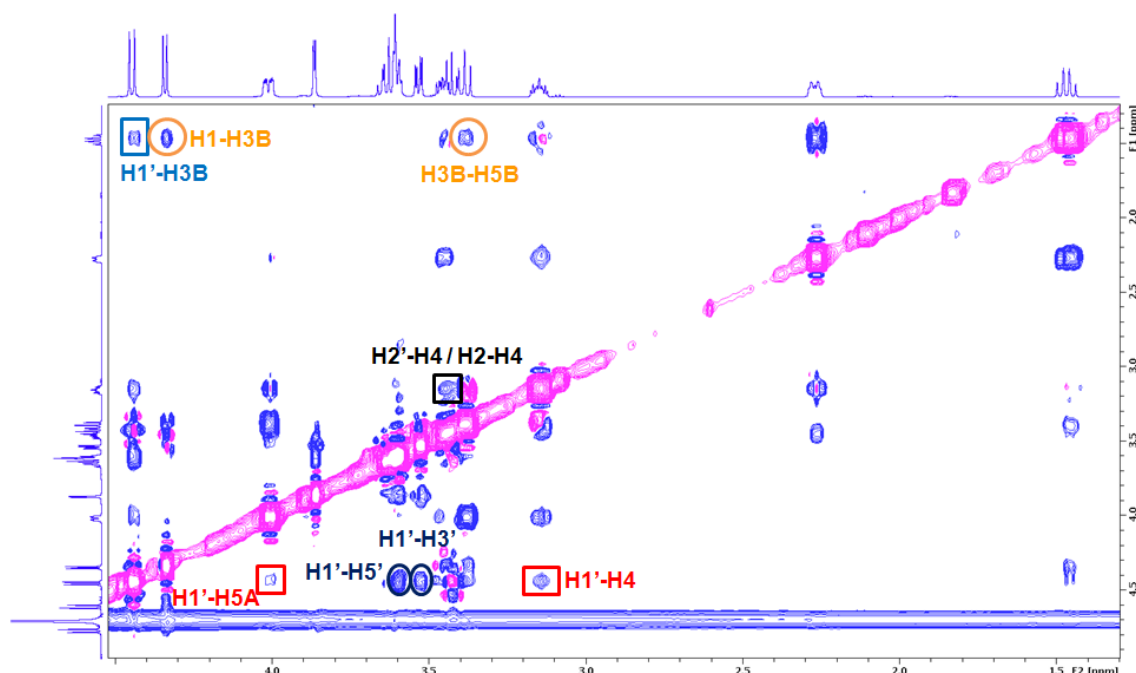


Figura 9: Espectro NOESY de **1** en estado libre. Las correlaciones intrarresiduales fueron representadas con círculos: *azules*, picos NOE característicos de la silla 4C_1 del residuo de Gal; *naranja*, picos NOE característicos de la silla 4C_1 del residuo de pentopiranososa; en cambio, las correlaciones interresiduales fueron representadas con cuadrados: *rojos*, H1'-H4 y H1'-H5A ($\text{syn } \phi_{4C1}/\text{syn } \psi_{4C1}$); *azul*, H1'-H3B ($\text{syn } \phi_{4C1}/\text{anti } \psi_{4C1}$); *negro*, H2'-H4 o H2-H4 ($\text{anti } \phi_{4C1}/\text{syn } \psi_{4C1}$).

El comportamiento de **2** se analizó de manera análoga, pero se tienen que considerar las cinco conformaciones, porque ambas sillas del anillo de 3-desoxi-*threo*-pentopiranososa fueron confirmadas en estado libre. Las conformaciones $\text{syn } \phi_{4C1}/\text{syn } \psi_{4C1}$ y $\text{syn } \phi_{1C4}/\text{syn } \psi_{1C4}$ deben mostrar el pico de cruce NOE entre el protón de la galactosa H1' y el protón H4, que no puede considerarse exclusivo de ninguna de ellas. Entre estas dos geometrías, la $\text{syn } \phi_{4C1}/\text{syn } \psi_{4C1}$ pudo ser confirmada por la presencia del contacto exclusivo H1'-H5A, pero la $\text{syn } \phi_{1C4}/\text{syn } \psi_{1C4}$ no pudo ser excluida. Además, la disposición $\text{syn } \phi_{4C1}/\text{anti } \psi_{4C1}$ fue detectada por su contacto exclusivo, H1'-H3B. Por último, respecto a la existencia de las conformaciones $\text{anti } \phi_{4C1}/\text{syn } \psi_{4C1}$ y $\text{anti } \phi_{1C4}/\text{syn } \psi_{1C4}$ los resultados de NOE no eran concluyentes, ya que el contacto H2'-H4 es difícilmente discriminado por la proximidad de sus señales. Por lo tanto, el inhibidor **2** en disolución muestra un equilibrio conformacional del que, con los datos experimentales de efecto Overhauser nuclear, no es posible excluir ninguna de las 5 geometrías distintas (Figura 10).

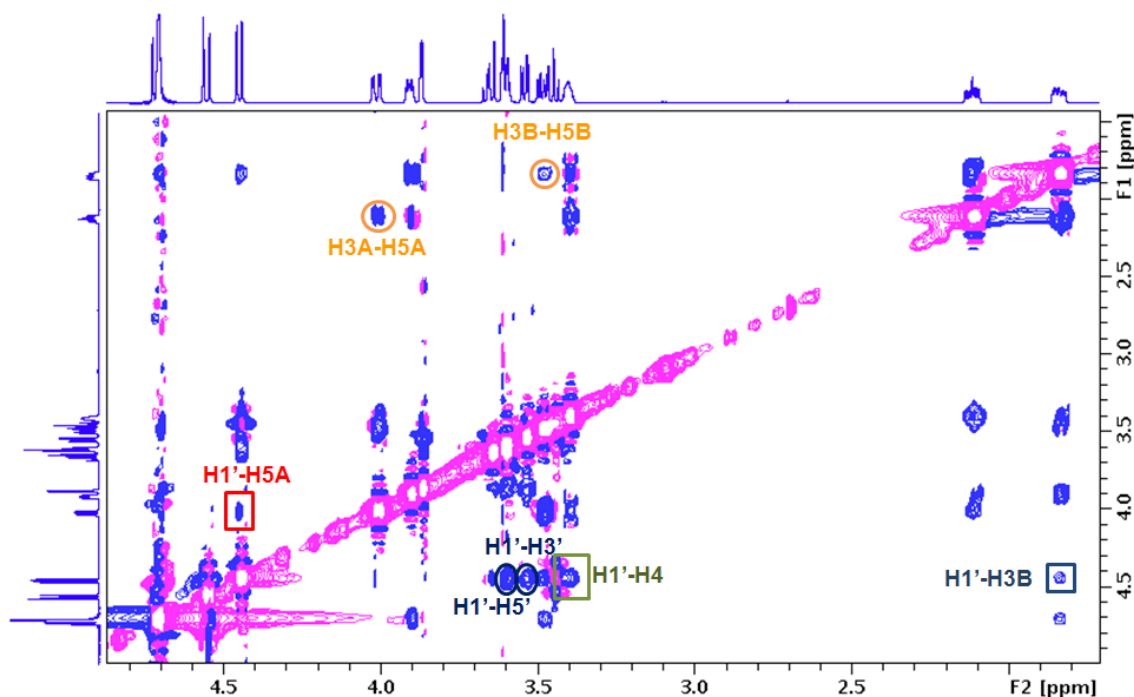


Figura 10: Espectro NOESY de **2** en estado libre. Las correlaciones intrarresiduales fueron representadas con círculos: *azules*, picos NOE característicos de la silla 4C_1 del residuo de Gal; *naranja*, picos NOE en el residuo de pentopiranos (ambas sillas representadas); en cambio, las correlaciones interresiduales fueron representadas con cuadrados: *verde*, H1'-H4 (syn ϕ_{4C1} /syn ψ_{4C1} y syn ϕ_{1C4} /syn ψ_{1C4}); *rojo*, H1'-H5A (syn ϕ_{4C1} /syn ψ_{4C1}); *azul*, H1'-H3B (syn ϕ_{4C1} /anti ψ_{4C1}).

5.2. Inhibidores de β -galactosidasa en estado asociado

La interacción de estos dos ligandos con la enzima β -galactosidasa de *Escherichia coli* fue estudiada experimentalmente mediante dos tipos de experimentos, STD (*Saturation Transfer Difference*)¹³⁶ y TR-NOESY (*transferred-NOESY*)¹⁴⁰. Los experimentos STD facilitaron información sobre el epítipo de reconocimiento y los TR-NOESY facilitaron información sobre la conformación que adoptan estos análogos de sustrato en el sitio de unión.

Los perfiles de los porcentajes de STD no fueron los mismos para **1** y **2**, figura 11, por lo tanto su epítipo de unión a esta proteína debe ser distinto. El mayor porcentaje de STD en el caso de **1** se observó para el protón Gal H4', junto a los protones aromáticos del extremo aglicónico, mientras que Gal H2' y H3' de este mismo residuo recibieron menos saturación que los protones del anillo de pentopiranos. La mayor saturación en **2** se observó en el anillo aromático, seguida de Gal H2'. Estos datos se compararon con los datos de STD de otro ligando de esta enzima de tipo tiogalactósido, IPTG (*IsoPropyl 1-ThioGalactopyranoside*). La comparación del patrón de %STD del residuo de Gal de este compuesto se asemeja más al perfil de reconocimiento del residuo Gal en **2** que en

1. El protón con mayor transferencia de saturación en el residuo de la Galactosa es el protón H2' tanto en el inhibidor **2** como en el IPTG.

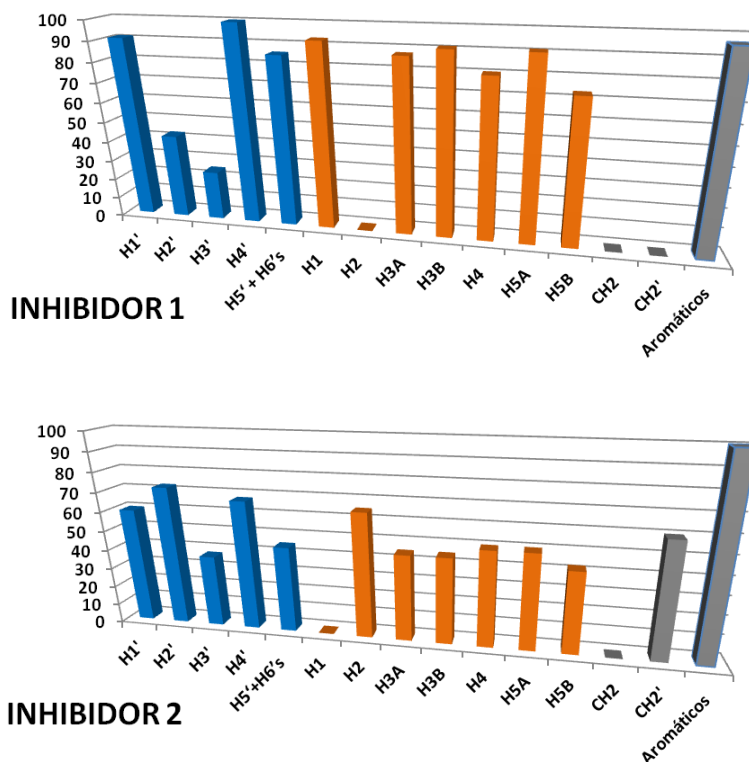


Figura 11: Porcentajes relativos de saturación de STD (2 segundos) para los inhibidores 1 y 2 (normalizados al protón que recibe mayor saturación). Las *barras azules* muestran los porcentajes del residuo de residuo de Galactosa, las *barras naranjas* los valores para el residuo de pentopiranososa y las *barras grises* para el aglicón bencilo. El 0% significa que no se ha podido medir experimentalmente el valor para esos protones.

Estos ensayos de STD se realizaron con una relación molar entre la concentración de ligando y la concentración de proteína (*ligando:proteína*) de 100:1. Estas condiciones, favorables para los experimentos de saturación, no corresponden con las óptimas del ensayo TR-NOESY. Los espectros TR-NOESY se adquirieron con muestras cuya relación *ligando:proteína* fue de 30:1. Los picos de cruce fueron de signo negativo (del mismo signo que la diagonal), diferentes de los existentes en ausencia del enzima (picos de cruce positivos o de signo opuesto a la diagonal). Este cambio de signo en los picos de cruce en el espectro NOESY es diagnóstico de un cambio en el tiempo de correlación del ligando (aprox. 0,3 ns) que se produce precisamente por unirse a la enzima adquiriendo las propiedades de ésta, cuyo tiempo de correlación es varios órdenes de magnitud más lento, como corresponde a un tetrámero de aproximadamente 500 kDa (aprox. 300 ns)²⁸⁶. Asimismo, los picos TR-NOESY contienen información sobre las geometrías de los ligandos en el sitio de unión.

En el espectro TR-NOESY del inhibidor **1** (Figura 12) se detectó la presencia de ambas sillas en el residuo de pentopiranos. La silla 4C_1 fue identificada mediante los NOEs H1'-H5B y H3B-H5B, mientras que la silla 1C_4 , muy poco poblada en el estado libre (<10%), fue detectada mediante un contacto de baja intensidad, H3A-H5A. En cuanto a las conformaciones alrededor del enlace glicosídico, la conformación $\text{syn } \phi_{4C1}/\text{anti } \psi_{4C1}$ no parece darse en el estado unido ya que su correspondiente NOE exclusivo, NOE H1'-H3B, claramente observable en el estado libre (Figura 9) está ausente en presencia de la enzima (Figura 12). La detección del NOE H1'-H4 y la ausencia del contacto NOE H1'-H5A corroboran la existencia de una conformación $\text{syn } \phi/\text{syn } \psi$, con la silla 1C_4 , pero no la misma conformación $\text{syn } \phi/\text{syn } \psi$, con la silla en disposición 4C_1 . Como ocurría en estado libre, el solapamiento entre los protones H2' y H2 no permitió ratificar la presencia de conformaciones $\text{anti } \phi/\text{syn } \psi$. Sin embargo, la presencia de la silla 4C_1 implica que la conformación $\text{anti } \phi_{4C1}/\text{syn } \psi_{4C1}$ debe existir en el estado unido. Por lo tanto, la disposición mayoritaria del inhibidor **1** en estado libre, $\text{syn } \phi_{4C1}/\text{syn } \psi_{4C1}$, no es reconocida por el enzima, que por el contrario une dos conformaciones minoritarias y de mayor energía en el estado libre, la $\text{anti } \phi_{4C1}/\text{syn } \psi_{4C1}$ y la $\text{syn } \phi_{1C4}/\text{syn } \psi_{1C4}$. Consecuentemente, el proceso de reconocimiento molecular debe incluir contactos intermoleculares estabilizantes que contrarresten una diferencia energética de al menos 1.5 kcal/mol.

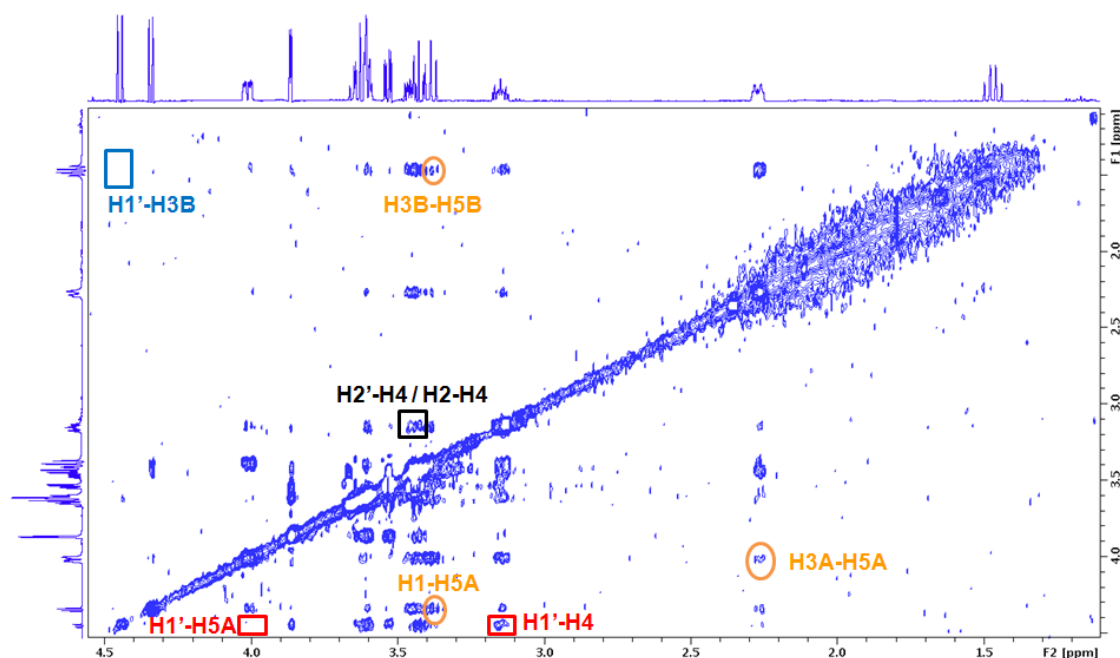


Figura 12: Espectro NOESY (200 milisegundos) del inhibidor **1** en presencia de la proteína (TR-NOESY). Las correlaciones intrarresiduales fueron representadas con círculos: *naranjas*, picos NOE característicos de la silla 4C_1 y de la silla 1C_4 del residuo de pentopiranos; en cambio, las correlaciones interresiduales fueron representadas con cuadrados: *rojos*, NOE H1'-H4 y ausente NOE H1'-H5A ($\text{syn } \phi_{1C4}/\text{syn } \psi_{1C4}$); *azul*, ausente NOE H1'-H3B ($\text{syn } \phi_{4C1}/\text{anti } \psi_{4C1}$); *negro*, NOE H2'-H4 ($\text{anti } \phi_{4C1}/\text{syn } \psi_{4C1}$) o NOE H2-H4.

El análisis de las conformaciones unidas del inhibidor **2** a la enzima (Figura 13) permitió concluir que eran las mismas que las detectadas para **1**, $\text{anti}\phi_{4C1}/\text{syn}\psi_{4C1}$ y $\text{syn}\phi_{1C4}/\text{syn}\psi_{1C4}$. Sin embargo, el proceso de reconocimiento molecular es diferente en este caso, ya que estas dos geometrías ya estaban presentes en su estado libre. El comportamiento de ambos ligandos denota que hay una correlación entre la conformación del enlace glicosídico y la disposición de los sustituyentes del anillo reductor en el centro activo. La geometría de la silla 4C_1 está ligada a un cambio en el ángulo de torsión ϕ , de $\text{syn}\phi$ a $\text{anti}\phi$, mientras que la conformación de menor energía, $\text{syn}\phi/\text{syn}\psi$, se correlaciona con la presencia de silla 1C_4 . En cualquier caso, la conformación más poblada en estado libre, $\text{syn}\phi_{4C1}/\text{syn}\psi_{4C1}$ (ausencia del contacto NOE H1'-H5A) no es reconocida por el enzima. En todos los casos, el enzima altera la conformación mayoritaria presente en estado libre, bien mediante la rotación alrededor de ϕ ($\text{syn}\phi \rightarrow \text{anti}\phi$), o modificando la geometría de la silla del extremo reductor (${}^4C_1 \rightarrow {}^1C_4$). Resultados análogos se han observado para otros análogos de sustrato, como se ha mencionado en el apartado de introducción.

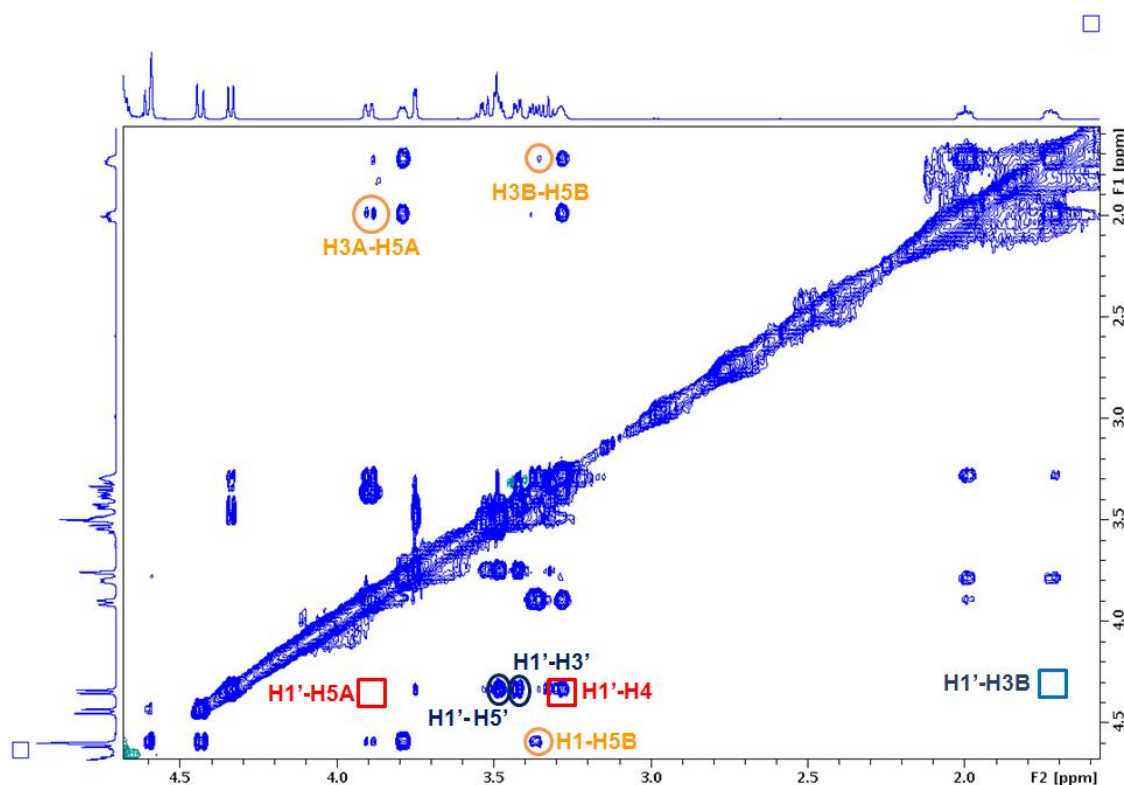


Figura 13: Espectro NOESY (200 milisegundos) del inhibidor **2** en presencia de la proteína (TR-NOESY). Las correlaciones intrarresiduales fueron representadas con círculos: *azules*, picos NOE característicos de la silla 4C_1 del residuo de Gal; *naranjas*, picos NOE característicos de la silla 4C_1 y de la silla 1C_4 del residuo de pentopiranos; *cuadrados rojos*, NOE H1'-H4 y ausente NOE H1'-H5A ($\text{syn}\phi_{1C4}/\text{syn}\psi_{1C4}$); en cambio, las correlaciones interresiduales fueron representadas con cuadrados: *rojos*, NOE H1'-H4 y ausente NOE H1'-H5A ($\text{syn}\phi_{1C4}/\text{syn}\psi_{1C4}$); *azul*, ausente NOE H1'-H3B ($\text{syn}\phi_{4C1}/\text{anti}\psi_{4C1}$).

5.3. Docking molecular de los inhibidores 1 y 2

Estos resultados experimentales se complementaron con el modelado de los complejos correspondientes. El análisis se desarrolló usando *docking* molecular mediante la herramienta *Glide* del programa *Schrödinger*. Los cálculos se realizaron empleando la estructura de rayos X de la β -galactosidasa de *E.coli* con el análogo de sustrato 2-fluoro-2-deoxy-D-lactosa, cuyo código PDB es 1JYY.²⁷⁸ Los estudios se hicieron por duplicado para ambos compuestos. Se utilizaron como punto de partida las dos geometrías del anillo de pentopiranos, 4C_1 y 1C_4 . Los resultados de los estudios de *docking* se agruparon en función de la disposición relativa en torno al enlace glicosídico y a la geometría de la silla del residuo aglicónico. Se usó como descriptor la distancia entre el protón anomérico de la Gal, H1', y el protón en posición 4 de la pentopiranos, H4. Esta distancia toma valores menores de 2.5 Å cuando el enlace es *syn* ϕ /*syn* ψ , mientras que, para cualquier disposición *anti*, la distancia es mayor de 3.1 Å. Para identificar la conformación de la silla del anillo de pentopiranos se empleó la distancia entre H3B y H5B. En la 4C_1 , la distancia entre estos dos protones es de 2.5 Å, mientras que si adopta la 1C_4 , los dos protones están separados por 4.1 Å.

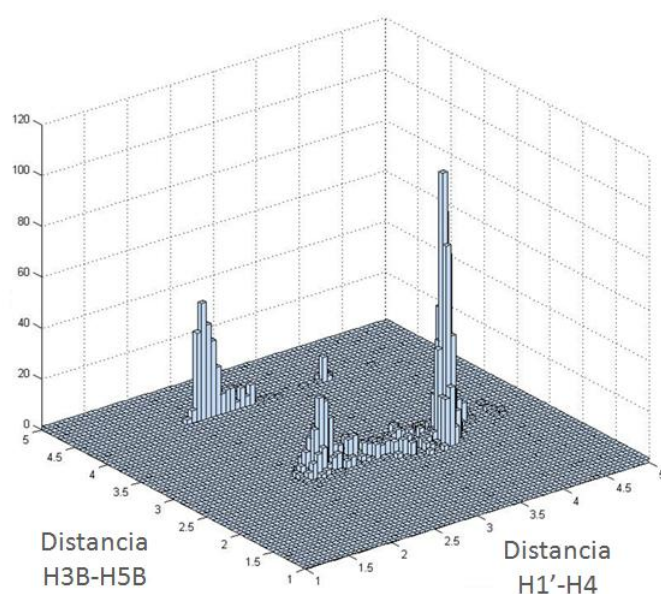


Figura 14: Histograma de población con las estructuras obtenidas en los estudios de *docking* del inhibidor **2** y la β -galactosidasa de *E.coli*. La distancia H3B-H5B es el descriptor de la conformación de la silla, mientras que la distancia H1'-H4 define la disposición relativa de los dos anillos de piranos.

Los cálculos de *docking* produjeron un número elevado de estructuras, que fueron agrupadas según sus distancias H1'-H4 y H3B-H5B. Para **2**, la representación indica que este compuesto puede adoptar tres conformaciones cuando interacciona con el enzima (Figura 14), las dos conformaciones más pobladas se corresponden con la

$\text{anti}\phi_{4C1}/\text{syn}\psi_{4C1}$ y la $\text{syn}\phi_{1C4}/\text{syn}\psi_{1C4}$ en concordancia con los resultados obtenidos experimentalmente en los estudios por RMN en presencia de la proteína (Figura 13). En cambio, los cálculos de docking del inhibidor **1** no están de acuerdo con las observaciones experimentales, ya que todas las estructuras de los cálculos de *docking* tienen una geometría 4C_1 en la pentopiranos. Estos cálculos no muestran la geometría $\text{syn}\phi_{1C4}/\text{syn}\psi_{1C4}$, pero la conformación más poblada coincide con la mayoritaria en el espectro TR-NOESY (Figura 12).

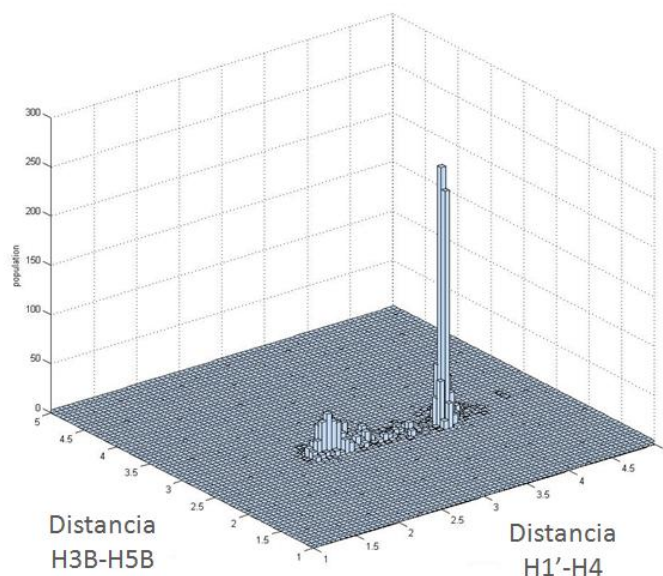


Figura 15: Histograma de población con las estructuras obtenidas en los estudios de *docking* del inhibidor **1** y la β -galactosidasa de *E.coli*. La distancia H3B-H5B es el descriptor de la conformación de la silla, mientras que la distancia H1'-H4 define la disposición relativa de los dos anillos de piranosa.

5.4. Dinámica molecular de los complejos proteína-carbohidrato

El análisis mediante simulaciones de dinámica molecular permitió determinar la estabilidad conformacional de los complejos. Como punto de partida se seleccionaron las conformaciones más pobladas, según los datos experimentales de RMN, $\text{syn}\phi_{1C4}/\text{syn}\psi_{1C4}$ y $\text{anti}\phi_{4C1}/\text{syn}\psi_{4C1}$. La estructura de la proteína se mantuvo estable durante el transcurso del cálculo (10 ns) y el ligando se mantuvo en el sitio de unión. De hecho, no se produjo ningún cambio ni fluctuación en la conformación de ambos anillos de piranosa; no hubo transición entre las 2 posibles sillas de la pentopiranos cuando la simulación se inició con cualquiera de las geometrías (Figura 16). Durante la simulación, también los ángulos de torsión alrededor del enlace glicosídico se mantuvieron estables. La variaciones de ϕ fueron mínimas, mientras que ψ presentó leves fluctuaciones, aunque manteniéndose en una disposición *syn* (Figura 17).

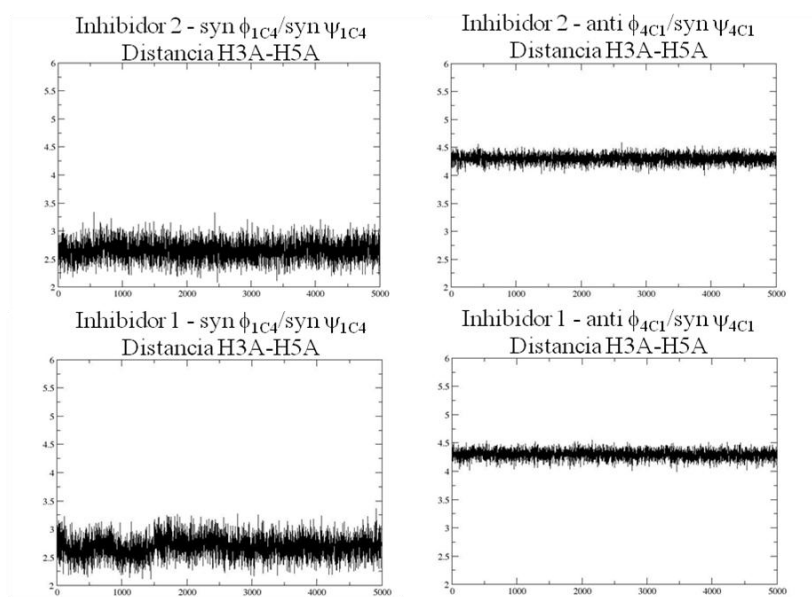


Figura 16: Distancias de los inhibidores **1** y **2** entre H3A y H5A de la pentopiranosita durante la simulación de DM. La conformación de la silla se mantiene a lo largo de toda la trayectoria.

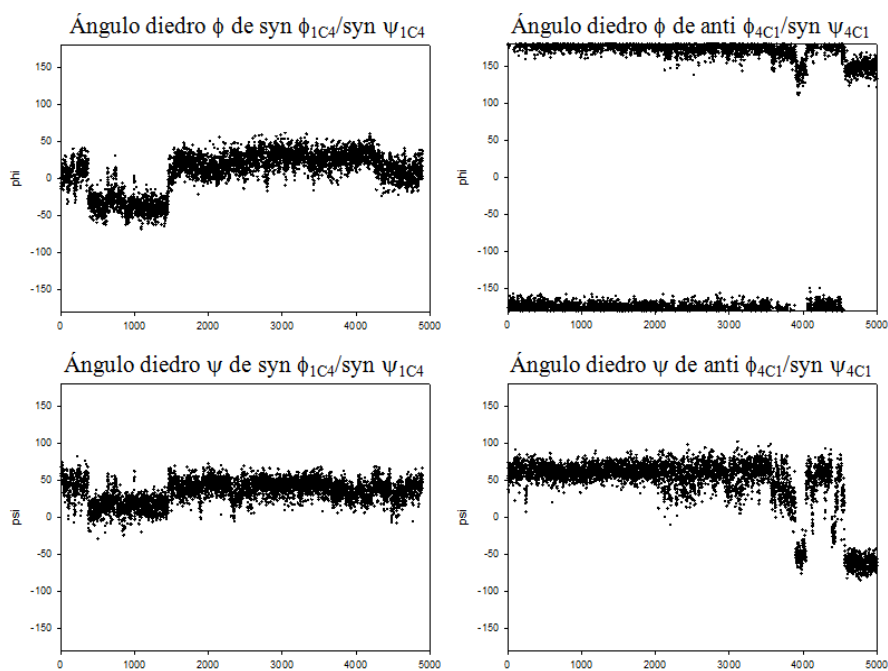


Figura 17: Valores de ϕ y ψ de **2** durante la simulación de dinámica molecular (10 ns, 5000 estructuras guardadas, cada 2 picosegundos). Las gráficas de la izquierda muestran la conformación $\text{syn } \phi_{1C4}/\text{syn } \psi_{1C4}$. ϕ se mantiene en conformación *syn*. Las gráficas de la derecha muestran la conformación $\text{anti } \phi_{4C1}/\text{syn } \psi_{4C1}$. ϕ se mantiene en conformación *anti*.

Además, se monitorizaron diversos contactos intermoleculares entre los ligandos y la β -galactosidasa. Es bien conocido que, en el reconocimiento molecular de carbohidratos, las interacciones CH- π tienen un papel esencial⁵⁵. En el caso concreto de este enzima, existen dos residuos aromáticos, Trp⁵⁶⁸ y Trp⁹⁹⁹ en el sitio de unión. Las estructuras de modelado molecular muestran que el residuo de galactosa establece interacciones con el Trp⁵⁶⁸. De hecho, la distancia entre el centroide del anillo de indol y el centroide del residuo de galactosa se mantuvo sin fluctuaciones, indicando la importancia de esta interacción estabilizante en ambos casos (Figura 18). Además, los complejos con ambos ligandos mostraron distintos enlaces de hidrógeno entre diferentes aminoácidos con el residuo de Galactosa: Gal O2 y Glu⁴⁶¹, Gal O3 y Glu⁵³⁷, Gal O4 y Asp²⁰¹, Gal O6 y Asn⁶⁰⁴, Gal O6 y Asp²⁰¹. Asimismo, se observaron interacciones transitorias entre el Gal O3 y Met⁵⁰² y entre Gal O2 y Tyr⁵⁰³.

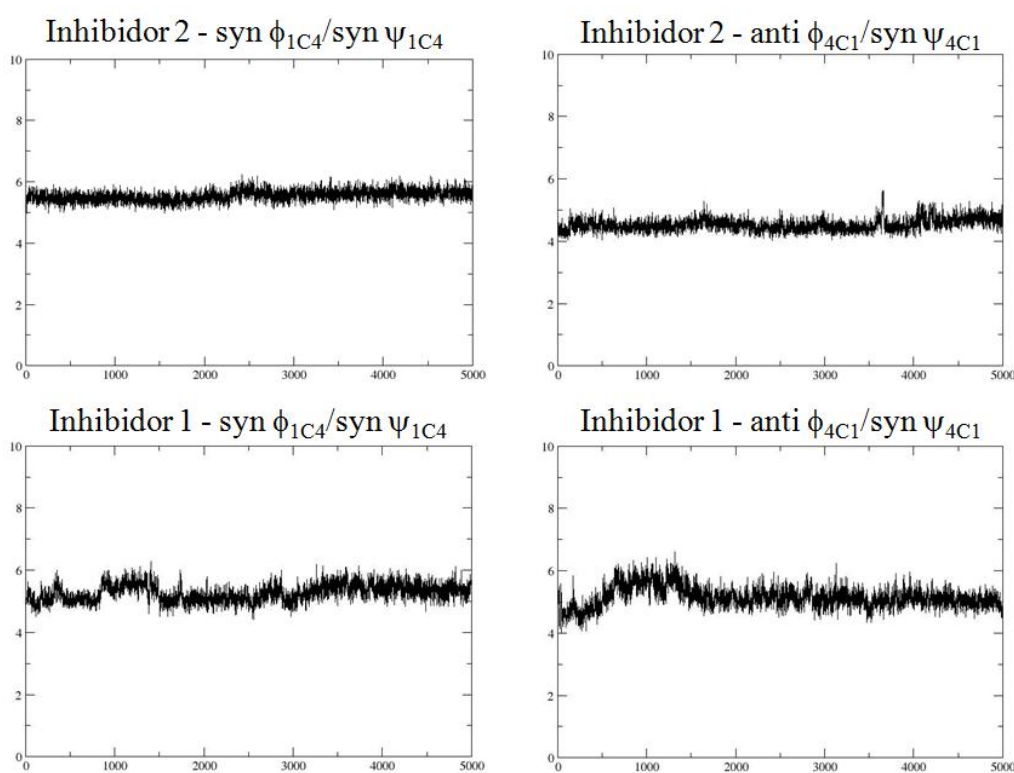


Figura 18: Distancia entre el centroide del anillo de indol de Trp⁵⁶⁸ y el centroide del residuo de galactosa durante la simulación de dinámica molecular. Ambos inhibidores muestran pocas fluctuaciones de esta distancia.

Estas interacciones colocan el residuo de galactosa en una posición muy definida en el sitio activo del enzima. Por tanto, el anillo de pentopiranososa se orientará de maneras muy diferentes dependiendo del valor de los ángulos de torsión alrededor del enlace glicosídico y establecerá distintas interacciones dependiendo de la silla que adopte. De hecho, en la conformación $\text{syn}\phi_{1C4}/\text{syn}\psi_{1C4}$, existe apilamiento entre este anillo y el Trp⁹⁹⁹. Sin embargo, en la $\text{anti}\phi_{4C1}/\text{syn}\psi_{4C1}$, no tiene lugar esta interacción. Debido a la

distinta configuración del hidroxilo en posición C2 entre **1** y **2**, existen también distintas interacciones implicadas en el reconocimiento. El inhibidor **2** muestra, en la conformación $\text{syn}\phi_{1C4}/\text{syn}\psi_{1C4}$, un enlace de hidrógeno entre su O2 ecuatorial con Asn^{102} . En cambio, en la conformación $\text{anti}\phi/\text{syn}\psi$, el O2 axial interacciona con Asn^{102} y con el NH de Val^{103} . Las dos geometrías son muy diferentes en el sitio de unión. El conformero $\text{anti}\phi_{4C1}/\text{syn}\psi_{4C1}$ muestra su anillo de pentopiranososa en una orientación que se asemeja a la que adoptaría el aglicón al final de la reacción enzimática. Para este inhibidor, orientaciones alternativas que dispusieran el O2 de la pentopiranososa orientado hacia el grupo indol del Trp^{999} , estarían desestabilizadas, ya que la interacción $\text{OH}-\pi$ está desfavorecida energéticamente ²⁸⁷.

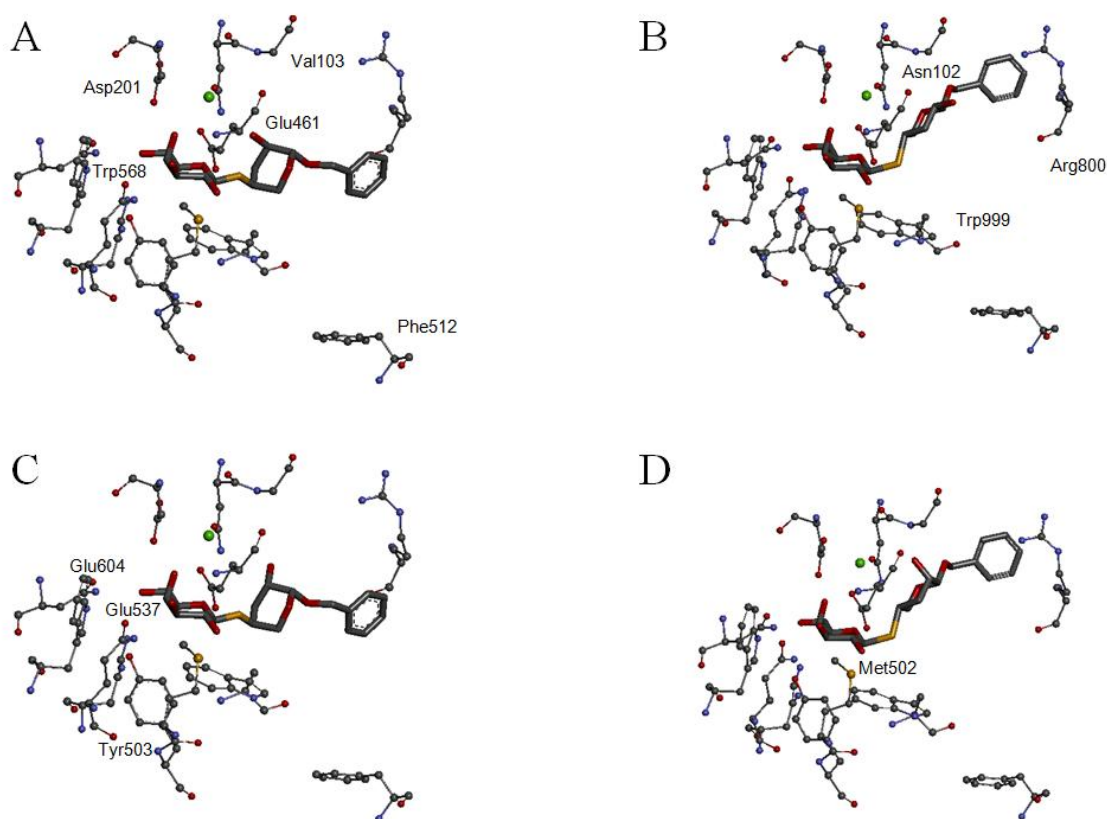


Figura 19: Representación esquemática de los complejos de **1** y **2** en el centro activo de β -galactosidasa de *E.coli*. A: **1** - $\text{syn}\phi_{1C4}/\text{syn}\psi_{1C4}$; B: **1** - $\text{anti}\phi_{4C1}/\text{syn}\psi_{4C1}$; C: **2** - $\text{syn}\phi_{1C4}/\text{syn}\psi_{1C4}$; D: **2** - $\text{anti}\phi_{4C1}/\text{syn}\psi_{4C1}$. Estas estructuras muestran la disposición de los dos conformeros de cada uno de los inhibidores al comienzo de la DM.

Todas las interacciones mencionadas anteriormente para **2** hacen referencia al inicio de la DM. Durante el transcurso de estos cálculos la distancia de la galactosa al Trp^{568} es constante, así que el extremo no-reductor se mantiene en la posición de partida. En

cambio, se han visto cambios en algunas interacciones y en la disposición del aglicón y del bencilo. Por ejemplo, el anillo aromático del grupo bencilo de estos compuestos también parece estar implicado en el reconocimiento molecular y su disposición en el centro activo. El muestreo de las distancias entre el anillo de bencilo del inhibidor **2** y dos fenilalaninas han demostrado que ambos conformeros de **2** muestran proximidad de este anillo aromático a la Phe⁵¹² al final de la simulación (Figura 20). De hecho, la Phe⁵¹² parece estar implicado en el reconocimiento de la conformación $\text{syn}\phi_{1C4}/\text{syn}\psi_{1C4}$. En cambio, cuando la conformación es $\text{anti}\phi_{4C1}/\text{syn}\psi_{4C1}$, el grupo bencilo establece una interacción catión- π con el grupo guanidinio de la Arg⁸⁰⁰ (Figura 21).

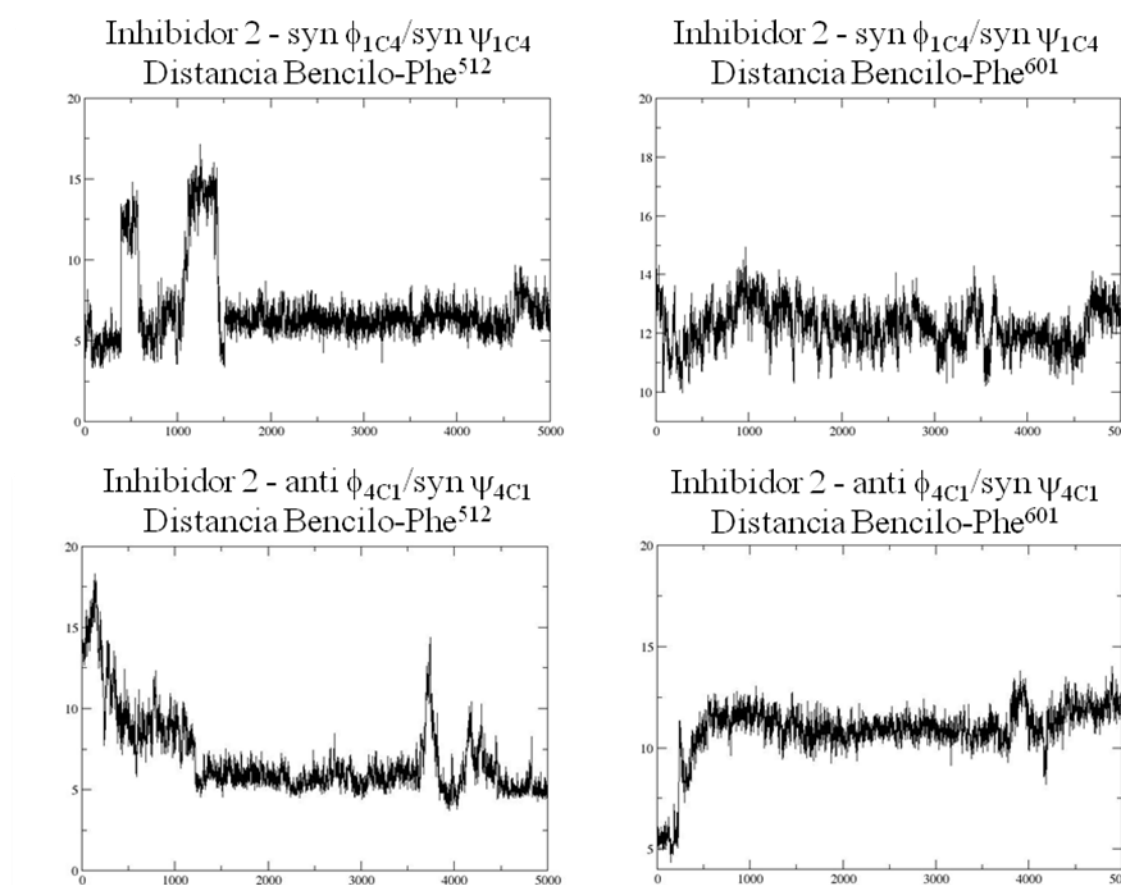


Figura 20: Distancias entre el centroide del anillo aromático del bencilo de **2** y el de los aminoácidos aromáticos Phe⁵¹² y Phe⁶⁰¹ durante la simulación de DM.

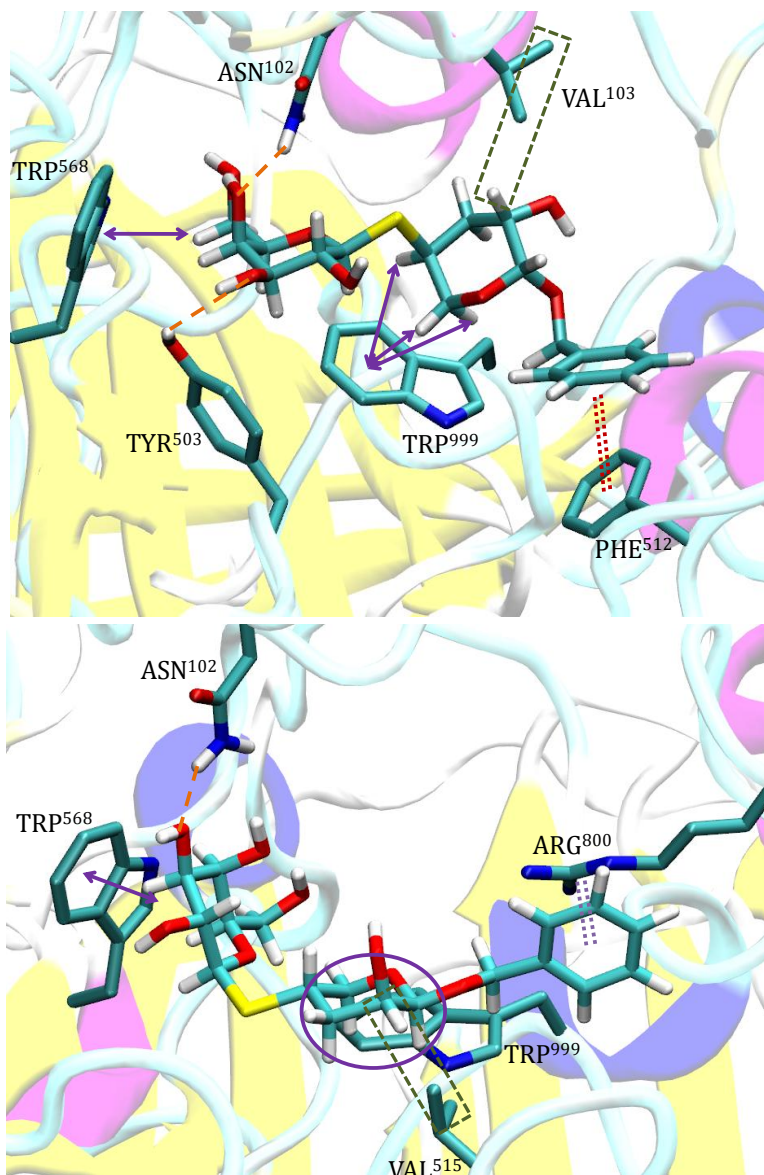


Figura 21: Estructura de **2** en el centro activo de la β -galactosidasa de *E.coli*. *Arriba*, Resultado de la DM con el conformero $\text{syn}\phi_{1C4}/\text{syn}\psi_{1C4}$. *Abajo*, Resultado de la DM con el conformero $\text{anti}\phi_{4C1}/\text{syn}\psi_{4C1}$.

El análisis de DM de las dos conformaciones de **1** con la β -galactosidasa sugiere que este compuesto interacciona de modo análogo. La Gal establece los mismos contactos que han sido mencionados anteriormente con el enzima. En la conformación $\text{syn}\phi_{1C4}/\text{syn}\psi_{1C4}$, también existen interacciones CH- π entre el residuo de pentopiranososa y Trp⁵⁶⁸, mientras que su O2 axial muestra un enlace de hidrógeno con Asn¹⁰². En la disposición $\text{anti}\phi_{4C1}/\text{syn}\psi_{4C1}$, la orientación ecuatorial del O2 no permite este contacto con la Asn¹⁰², que ahora forma un enlace de hidrógeno con el O5 del ciclo del pentopiranososa. Este menor número de interacciones entre **1** y la enzima permite explicar su menor inhibición frente a **2**.

Por otra parte, existe un cambio conformacional notable en el inhibidor **1** al pasar del estado libre, cuya conformaciones mayoritarias tienen el anillo de pentopiranososa en disposición 4C_1 , al unido, cuyas conformaciones no están prácticamente representadas en el estado libre, y que debe estar penalizado energéticamente. Sin embargo, para el inhibidor **2**, las conformaciones reconocidas ya estaban presentes en solución, por lo que la penalización entrópica debe ser menor.

Finalmente, el distinto patrón de STD del residuo de galactosa del inhibidor **1** (figura 7) sugiere que este compuesto debe adoptar otra pose en el centro activo de la enzima o, más probable, interaccionar adicionalmente con otra región de la proteína. De hecho, los ensayos de inhibición de este compuesto mostraron un comportamiento no-competitivo.

5.5. Conclusiones: Desarrollo de inhibidores de glicosidasas mediante modificaciones en el anillo aglicónico

Existen diversas aproximaciones al desarrollo de inhibidores de glicosidasas mayormente basadas en modificaciones sobre el propio residuo que sufre la ruptura del enlace glicosídico en el presente trabajo las diferencias observadas en la actividad inhibidora de **1** y **2** frente a la β -galactosidasa de *E.coli* muestran la posibilidad de desarrollar inhibidores mediante la modificación en el residuo aglicónico. El inhibidor **1** sólo se diferencia de **2** en la configuración de uno de los hidroxilos del anillo de pentopiranososa. Sin embargo, su constante de inhibición es 800 μ M mientras que la de **2** es 32 μ M. Ambos compuestos de acuerdo con los datos de RMN y apoyados por técnicas de modelado molecular pueden adoptar dos conformaciones diferentes en el estado unido, La disposición anti ϕ /syn ψ tiene la piranosa del extremo reductor en conformación 4C_1 , mientras que la geometría syn ϕ /syn ψ muestra el residuo de pentosa en conformación 1C_4 . El inhibidor **2** ya presentaba estas conformaciones en su estado libre, mientras que **1** sólo adoptaba una disposición de silla 4C_1 en el anillo de pentopiranososa. Por lo tanto, el enzima modifica y selecciona la conformación reconocida. El inhibidor más potente, **2**, se une al centro activo de la enzima preferentemente con el anillo aglicónico en conformación 1C_4 ; mientras que **1** muestra una contribución similar de ambas conformaciones. La existencia de distorsión en el anillo de piranosa se ha descrito en diversas ocasiones, aunque siempre referido al residuo no reductor, sobre el que se produce la hidrólisis.²⁸⁸⁻²⁹⁰ De hecho, el estado de transición de la reacción de hidrólisis debe mostrar una piranosa distorsionada. Sin embargo, este estudio sugiere la posibilidad de encontrar inhibidores relativamente potentes mediante modificaciones en el anillo de piranosa del extremo reductor. La diferencia entre las constantes de inhibición en 25 veces demuestra que se puede modular la actividad de análogos de sustrato mediante las variaciones químicas y conformacionales en el anillo aglicónico.

5.6. Materiales y métodos

Los compuestos fueron sintetizados en el grupo del Dr. Oscar Varela del CIHIDECAR-CONICET de la Universidad de Buenos Aires.^{284,291}

Sus constantes de inhibición fueron medidas utilizando la misma enzima que se ha usado en el estudio de RMN, adquirida a Sigma-Aldrich. Fue necesario eliminar el tampón TRIS que contiene la muestra comercial y llevarla a un tampón en agua deuterada. La preparación se sometió a una etapa de intercambio de buffer mediante ultrafiltración con VIVASPIN 6 con un corte de peso molecular de 10 kDa. Se realizaron 10 pasos de dilución-concentración usando un tampón fosfato 20mM, a pH 7.2, con MgCl₂ 1mM, en D₂O para realizar los experimentos de RMN.

5.6.1. Resonancia Magnética Nuclear

Los experimentos de RMN se realizaron en un espectrómetro Bruker AV600 utilizando las secuencias de pulsos estándar del programa TOPSPIN. Para los experimentos NOESY y TR-NOESY se utilizó la secuencia *noesygppl19*, para los NOE selectivos se usó la secuencia *selnogp* y, para los experimentos STD, la secuencia *std2*.

Los espectros NOESY en estado libre se adquirieron con dos tiempos de mezcla, 200 y 500ms. Los experimentos TR-NOESY, en presencia de la proteína (relación molar proteína:ligando 1:30), se realizaron con 80 y 200ms. No fue necesaria la utilización de un filtro T2, o de *spin-lock*, para eliminar las señales de fondo de la proteína porque el gran tamaño de la β -galactosidasa de *E.coli* (tetrámero de 116.3 kDa) provoca que sus señales se ensanchen hasta confundirse con la línea base. Los experimentos STD se realizaron con una relación molar 1:100. La frecuencia off-resonance se centró a 100 ppm, mientras que la on-resonance se fijó a -0.5ppm, saturando la región de alifática de la proteína. Se realizaron experimentos STD con distintos tiempos de saturación (0.5, 1, 1.5, 2 y 2.5 segundos). Los datos se normalizaron al protón que recibió mayor saturación.

5.6.2. Cálculos de Mecánica Molecular

Los mapas energéticos²⁹² se calcularon mediante la herramienta *Coordinate Scan* dentro del paquete de programas *Maestro*¹⁸². Se seleccionó *AMBER** como campo de fuerzas y el método de cálculo utilizó un protocolo *PRCG*, que realizó una minimización energética cada vez que los ángulos diedros se modificaron sistemáticamente en 10°. Se calcularon mapas energéticos para ambas conformaciones de silla del residuo de pentopiranososa para registrar la influencia de este anillo en la disposición del enlace glicosídico. La representación de las distancias representativas entre los protones de estas moléculas permitió determinar los confórmeros en estado libre y en estado unido por comparación con los datos experimentales de NOE.

5.6.3. Docking molecular

La interacción del enzima con **1** y **2** se analizó mediante *docking* molecular utilizando el programa *Glide* (versión 5.5²⁹³) de *Schrödinger*. Se seleccionaron los parámetros estándar, siguiendo un método SP (*standard precision*), sin someter al ligando a ninguna restricción torsional ni conformacional. Los cálculos usaron como punto de partida la estructura de rayos X de la β -galactosidasa de *E.coli*, código 1JYY en el *Protein Data Bank*²⁷⁸. En primer lugar, se trató la estructura con la herramienta *Protein Preparation Wizard* de *Maestro*. Se eliminaron las moléculas de agua y los iones de la estructura que no estaban implicados en el proceso de reconocimiento, se añadieron los hidrógenos polares a la estructura y se revisó manualmente el estado de protonación de las histidinas. Finalmente, se eliminó el ligando que contenía la estructura, *2-F-lactosa*. Seguidamente, se utilizó la herramienta *LigPrep* para generar las conformaciones requeridas de **1** y **2** que se ensayaron con el programa *Glide*. Después del tratamiento tanto de la proteína y como del ligando, se realizó el *docking* molecular usando una caja cúbica de 25 Å de lado, centrada en el centro activo del enzima. Los resultados se analizaron mediante histogramas, tal y como se describe en el texto.

5.6.4. Dinámica Molecular

Las simulaciones de dinámica molecular se realizaron mediante el módulo *Sander* de AMBER10²⁹⁴, usando el campo de fuerzas AMBER con los parámetros ff99 y GLYCAM²⁹⁵. Los sistemas se neutralizaron con 32 contraiones de Na⁺ y, posteriormente, se solvataron usando agua de tipo TIP3P en un caja de solvente explícito cuyas dimensiones tuvieron una distancia mínima de 8 Å al complejo. Se aplicó el algoritmo Shake²⁹⁶, y las interacciones electrostáticas se trataron utilizando el método *particle mesh Ewald*²⁹⁷. En primer lugar, las estructuras de los complejos ligando:proteína se sometieron a un proceso de *annealing* durante 25 picosegundos desde una temperatura de 100 K a una de 300 K, y seguidamente se minimizaron con las moléculas de soluto completamente restringidas. Las restricciones se redujeron paulatinamente hasta permitir la libertad de todos los elementos del sistema. Posteriormente, se realizó una fase de calentamiento durante 20 ps entre 100 K y 300 K sin ninguna restricción en los componentes del sistema. Tras estas fases de preparación, se realizó la simulación de dinámica molecular propiamente dicha, con una duración de 10 nanosegundos en condiciones isotérmicas e isobáricas. Las coordenadas fueron recopiladas cada 2 ps.

Las trayectorias obtenidas de las dinámicas se analizaron mediante los programas AMBER y VMD²⁹⁸. La desviación cuadrática media (RMSD) del esqueleto proteico se calculó mediante el módulo *ptraj* de AMBER. Además, se midieron distancias entre parejas de protones seleccionadas y los ángulos de torsión glicosídicos durante la simulación. La conformación de los inhibidores se caracterizó mediante el valor de ϕ y ψ , y las distancias entre los protones del residuo de pentopiranos. Además, se midieron otras distancias intermoleculares para definir contactos entre la proteína y el ligando, especialmente enlaces de hidrógeno e interacciones de apilamiento.

Discusión

Los trabajos descritos en esta Memoria tienen en común la metodología empleada: Aplicaciones de la espectroscopia de RMN al estudio de la estructura y conformación de biomoléculas y al estudio de procesos de reconocimiento molecular.

En todos los capítulos se han empleado diferentes métodos y se han evaluado distintos parámetros de RMN con información clave de estructura y conformación: desplazamientos químicos, parámetros de relajación, constantes de acoplamiento y efectos *Overhauser* nuclear. Todos estos parámetros se han empleado como denominador común para acceder a la información estructural, conformacional, dinámica y de interacción de las distintas moléculas empleadas.

La Tesis consta de dos partes bien diferenciadas, con objetivos diferentes y con entidad en sí mismas: el estudio estructural de péptidos y de glicanos. En los distintos capítulos estas moléculas se han combinado, y se ha evaluado, por ejemplo, el comportamiento conformacional de glicopéptidos o se ha estudiado la interacción de carbohidratos, naturales y sintéticos, con distintas proteínas.

Así, en la primera parte, se ha abordado el estudio de la estructura de péptidos relacionados con el dolor, prestando especial atención a las variaciones de desplazamiento químico, constantes de acoplamiento y NOEs medidos para las distintas moléculas.

El estudio de RMN de análogos del péptido opiorfina ha mostrado diferencias conformacionales en los análogos que se ha sustituido el residuo L-Phe en posición 3 por su enantiómero D-Phe, que están relacionadas con su actividad biológica. Los péptidos con un residuo D-Phe muestran una preferencia conformacional definida en su ángulo χ_1 , sin embargo los péptidos que contienen el residuo natural L-Phe presentan una mayor flexibilidad.

En el capítulo de los análogos glicosilados de la nociceptina, además de estudiar detenidamente las variaciones de desplazamiento químico y constantes de acoplamiento, se han empleado fundamentalmente experimentos NOESY asistidos por métodos computacionales (CYANA). Los datos experimentales confirman que el glicopéptido más activo tiene mayor tendencia a formar α -hélice que el péptido natural. Por lo tanto, el papel del glicano ha sido caracterizado mediante la combinación de RMN y modelado molecular.

El estudio de la interacción de los quitoooligosacáridos etiquetados con grupos difluoroacetamida en los residuos de GlcNAc se ha basado en las variaciones de la relajación transversal de los núcleos ^1H y ^{19}F del ligando al interaccionar con la

proteína. De hecho, el ensanchamiento diferencial observado en las distintas señales ha permitido deducir el epítipo de estos ligandos al ser reconocidos por WGA.

Por último, en el estudio de la inhibición de la β -galactosidasa por distintos glicomiméticos se ha utilizado, de nuevo, la combinación de distintos métodos de RMN (desplazamientos químicos, parámetros de relajación, constantes de acoplamiento y efectos Overhauser nucleares). El análisis exhaustivo de estos parámetros, medidos para los inhibidores tanto en estado su estado libre como asociados a la enzima, ha permitido caracterizar su comportamiento dinámico en las distintas condiciones. En este capítulo también se han usado métodos computacionales de manera exhaustiva.

En definitiva, el trabajo realizado en esta Tesis me ha permitido adquirir una formación avanzada en distintos métodos y técnicas de RMN y en el empleo de métodos computacionales para apoyar la interpretación de los datos experimentales. He aprendido a preparar las muestras para la realización de los experimentos de RMN en distintas condiciones experimentales (ligandos libres y en estado asociado, en medios tamponados y micelas). Además, he aprendido a preparar, calibrar, ejecutar e interpretar gran variedad de experimentos de RMN. Me he formado en los métodos necesarios en la asignación de espectros de péptidos y glicopéptidos mediante métodos de correlación homonucleares, y en el empleo de programas de generación de estructuras.

En la segunda parte de la Tesis, he profundizado en el estudio de procesos de reconocimiento molecular de las interacciones carbohidrato-proteína, empleando dos sistemas diferentes de reconocimiento, una lectina y un enzima. La metodología usada es complementaria a la empleada en la primera parte. Además, he aprendido los métodos de RMN que se aplican para estudiar estas interacciones ligando-receptor desde la perspectiva del primero, aplicando experimentos clave como el STD y el TR-NOESY. Conjuntamente se ha desarrollado una metodología que emplearía unas sondas de RMN para detectar interacciones, basadas en etiquetas de ^{19}F . El método es de interés general y puede aplicarse a otros sistemas. Este capítulo me ha permitido formarme en RMN de ^{19}F , como complemento a la RMN de ^1H , para conseguir una visión global del proceso de reconocimiento molecular.

Finalmente, he estudiado la interacción de inhibidores, análogos del sustrato, con un enzima mediante RMN. Esta investigación me ha facilitado una formación en el estudio conformacional de esta clase de compuestos mediante la combinación del análisis de constantes de acoplamiento y NOEs, apoyados por métodos computacionales para conseguir información sobre la geometría de ligandos complejos. Además, he adquirido conocimiento sobre cómo conseguir información sobre la conformación de ligandos cuando están asociados a un receptor. En este capítulo, de nuevo, se han usado métodos experimentales de RMN apoyados por métodos de modelado molecular.

Es decir, la Tesis Doctoral me ha permitido conseguir una visión global del uso de la RMN al estudio de procesos de reconocimiento molecular en química y biología. He

usado muy diversas técnicas, métodos y protocolos. He aprendido a asignar espectros de péptidos y carbohidratos, a generar estructuras de los mismos basados en los datos experimentales apoyados en cálculos teóricos, y he aprendido a estudiar procesos de reconocimiento molecular mediante RMN, también apoyado en otros métodos computacionales.

Por último, esta formación privilegiada me ha permitido encontrar un trabajo en investigación dentro de una empresa multinacional farmacéutica, en la que se emplean todos estos métodos al máximo nivel.

Conclusiones

En esta tesis doctoral se ha abordado el estudio de cuatro sistemas diferentes mediante técnicas de RMN, que han sido complementadas con un análisis de mecánica y dinámica molecular que facilitan la comprensión estructural de estos sistemas:

- Estudio estructural de análogos del péptido opiorfina

En primer lugar, el *barrido de alanina* en la opiorfina demostró la importancia del residuo de su posición 3, Phe³, en la unión de este péptido a las enzimas NEP y APN.

Posteriormente, diversas modificaciones en el residuo aromático mostraron principalmente análogos que disminuían su actividad frente a ambas Zn-metalopeptidasas, pero los dos análogos que sustituían este aminoácido natural por su enantiómero, D-Phe, mejoraban su inhibición frente a APN y reducían drásticamente la actividad frente a NEP.

La opiorfina y 7 análogos en la Phe³ fueron asignados por RMN, y se analizaron las diferencias espectroscópicas que indican cambios estructurales en disolución.

La conformación de los análogos que contienen D-Phe es más restringida que la del resto de péptidos, determinando que estos péptidos tengan una mejora en su actividad frente a APN.

Los resultados derivados de estos estudios han sido publicados en *Journal of Medicinal Chemistry* (DOI: 10.1021/jm2012112). Se ha identificado una estrategia para mejorar la potencia frente a APN mediante análogos con modificaciones en la posición 3, que son metabólicamente más estables que el péptido natural.

- Estudio estructural de análogos glicosilados de nociceptina

Los cuatro péptidos analizados, la nociceptina y sus análogos glicosilados, tienen distintos valores de actividad frente a los receptores de membrana NOP (drNOP). La [Ser¹⁰-O- α -D-GalNAc]-nociceptina es el glicopéptido de mayor afinidad y la [Thr⁵-O- α -D-GalNAc]-nociceptina es el glicopéptido de menor afinidad.

La nociceptina y 3 análogos glicosilados de este péptido fueron asignados por RMN, y se analizaron las diferencias estructurales en disolución acuosa y en presencia de micelas de SDS.

En disolución no hay diferencias significativas entre ellos, ninguno de estos péptidos adopta una estructura secundaria definida como demostraron tanto el análisis de RMN como de DC.

En cambio, estos péptidos han mostrado diferentes comportamientos en un medio mimético de membrana como las micelas de SDS.

Los datos estructurales muestran una mayor tendencia de α -hélice para la [Ser¹⁰-O- α -D-GalNAc]-nociceptina, que tiene la mayor actividad biológica. Esta correlación entre el porcentaje de α -hélice y actividad ha sido también propuesta en otros trabajos.

De manera análoga, se ha encontrado que la menor tendencia a formar α -hélice de los otros glicopéptidos correlaciona con un descenso de actividad. La [Thr⁵-O- α -D-GalNAc]-nociceptina que tiene preferencia por una geometría plegada con menor porcentaje de α -hélice, es el análogo que presenta la afinidad más baja por los receptores drNOP.

Los resultados derivados de estos estudios han sido publicados en *Organic and Biomolecular Chemistry* (DOI: 10.1039/C1OB05197K). Se ha puesto de manifiesto la relación estructura-actividad en la interacción entre la nociceptina y sus receptores de membrana.

- El uso de etiquetas difluoroacetamida como sonda para detectar interacciones carbohidrato-proteína

Se ha realizado la asignación de las señales de RMN de ¹H y ¹⁹F de los análogos de N,N'-diacetil quitobiosa y N,N',N''-triacetil quitotriosa con etiquetas difluoroacetamida.

La interacción de estos dos análogos frente a WGA ha sido analizada tanto por experimentos de ¹H como de ¹⁹F RMN.

Los cálculos *ab initio* muestran que los dos flúores geminales mejoran levemente la interacción CH/ π , ya que polarizan el enlace C-H contiguo y determinan que exista una orientación favorecida energéticamente.

Los cálculos de dinámica molecular de los análogos de quitobiosa y de quitotriosa asociados a un dominio heveína están de acuerdo con los resultados de los experimentos de RMN. El análogo de quitobiosa interacciona principalmente a través de su extremo no-reductor. En cambio, el análogo de quitotriosa puede adoptar dos modos de unión que determinan que las señales del residuo del extremo no-reductor y el residuo central sean más fuertemente afectadas que las señales del extremo reductor.

Los resultados derivados de estos estudios han sido publicados en *Chemistry - A European Journal* (DOI: 10.1002/chem.201501248). Se muestra la utilidad de esta estrategia para el diseño de ligandos en procesos de reconocimiento molecular en los que estén implicados residuos GlcNAc, GalNAc o NeuNAc.

- Inhibidores de β -galactosidasa a través de modificaciones en anillo aglicónico

En este estudio se analizan dos inhibidores que se diferencian únicamente en la configuración de uno de los hidroxilos del anillo aglicónico, pero la diferencia entre sus constantes de inhibición es 25 veces (Inhibidor **1**, 800 μ M; e Inhibidor **2**, 32 μ M).

A partir de los datos de constantes de acoplamiento y de los contactos NOE de ambos inhibidores, se mostró que estos compuestos no se comportan del mismo modo en estado acuoso. **1** sólo presenta conformaciones cuyo anillo de pentopiranososa adopta una silla 4C_1 . En cambio, **2** presenta conformaciones que adoptan tanto una silla 4C_1 como una silla 1C_4 .

Los datos de RMN de TR-NOESY en presencia de la β -galactosidasa, apoyados por técnicas de modelado molecular, muestran que ambos compuestos pueden adoptar dos conformaciones en estado unido. La disposición anti ϕ /syn ψ que tiene la pentopiranososa del aglicón en conformación 4C_1 , y la geometría syn ϕ / syn ψ cuyo residuo de pentopiranososa está en conformación 1C_4 .

El inhibidor **2** ya presentaba ambas conformaciones en su estado libre, mientras que **1** principalmente adoptaba una disposición de silla 4C_1 en el anillo de pentopiranososa. La enzima debe seleccionar o modificar las conformaciones en el proceso de reconocimiento. La existencia de distorsión en anillos de piranososa se ha descrito en diversas ocasiones, aunque siempre referido al residuo no reductor, sobre el que se produce la hidrólisis. En la presente tesis se muestran modificaciones en el anillo de piranososa del residuo aglicónico. El inhibidor más potente, **2**, se une al centro activo de la enzima especialmente con el anillo aglicónico en conformación 1C_4 ; mientras que **1** muestra una contribución similar de ambas conformaciones.

Los resultados derivados de estos estudios han sido publicados en *Chemistry - A European Journal* (DOI: 10.1002/chem.201203673). Se muestra la capacidad de modular la actividad inhibitoria de análogos de sustrato mediante las variaciones químicas y conformacionales entorno al anillo aglicónico.

Summary

Introduction

Molecular recognition is involved in numerous life processes. The exploration of the ability to control intermolecular interactions has become a new frontier in chemistry, biochemistry, and cellular and molecular biology. Thus, understanding the variables which are involved in the interactions between biological targets and their ligands in these life processes has the potential to rationalize binding and selectivity features, and therefore to assist in the design of new molecular probes and of novel therapeutic agents. In fact, the detailed study of 3D structures has led to a deeper comprehension of the molecular basis of drug–receptor recognition, and thus to enhance affinity of inhibitors toward their targets.

Despite the fact that X-ray crystallography is a powerful tool, which provides structural information of both small and large molecules, the technique is limited in cases where the molecules are difficult to crystallize. Besides, it offers a static view of the binding process. In fact, and specially in the case of small molecules, both the ligand and/or the receptor might adopt several conformations in solution, information not accessible by X-ray analysis, that can simplify and misunderstand these biological processes by partial or incomplete piece of information¹²⁰⁻¹²¹.

A combined approach of experimental Nuclear Magnetic Resonance (NMR) and theoretical molecular mechanics and dynamics could clarify the system dynamics and provide structural information which couldn't be easily accessed with other techniques⁴⁻⁶. In this thesis the analysis of four different systems of biological relevance has been accomplished by this combined approach. Each system will be discussed in next chapters.

Chapter 1 – Structural analysis of opiorphin analogs by NMR

Opiorphin (QRFSR) is a peptide which slows down enkephalin degradation, and consequently has antinociceptive properties^{157,299}. This peptide, which was isolated from human saliva, inhibits both neprilysin (NEP) and aminopeptidase N (APN) ectoenkephalinases, modulating the degradation of enkephalins by its interaction with these membrane-bound zinc metalloenzymes. Its antinociceptive effects have been demonstrated in some rodent models¹⁵⁹⁻¹⁶⁰. Thus, analgesic processes can be modulated by a natural peptide by affecting the catabolism of enkephalins¹⁵⁶. In this first chapter different opiorphin analogs were characterized by their activity against *h*NEP and *h*APN. The most remarkable findings were analyzed by NMR conformational studies in solution and molecular mechanic calculations in order to understand the changes in their potency.

The activity of several opiorphin analogs were measured *versus* hNEP-Ectopeptidase and hAPN-ectopeptidase by Fluorimetric Assays¹⁷⁸⁻¹⁷⁹. Alanine scanning in opiorphin sequence demonstrated that most Ala-substituted analogs were inferior inhibitors *versus* both targets, and the most drastic effect was observed in Ala³ substitution. Thus, the aromatic residue, Phe³, should play a relevant role in the recognition of opiorphin by hNEP and hAPN. Furthermore, different modifications at this aromatic residue showed a drastic effect in the dual inhibition. The most remarkably finding was that the D-Phe³-containing peptides were one order of magnitude more potent than the parent peptide in APN assays, however inactive in NEP assays (IC₅₀>100μM).

These activity data were rationalized from a structural perspective by solution NMR experiments by comparing the parent peptide with its analogs. NMR corroborates the different tendencies between L-Phe³-containing and D-Phe³-containing peptides, while no significant differences were appreciated for the fluorinated derivatives. Chemical Shift Indexes (CSI) for the H α protons showed that the D-Phe³-containing peptides diminish their β -sheet-forming propensity at position 2. Besides, the vicinal coupling constant $^3J_{\text{NH-H}\alpha}$ for Arg² changed slightly, showing lower values for the D-Phe³ analogs. However, the most substantial difference between these peptides was the chemical shifts of Arg² H γ 2 and H γ 3 protons, being both H γ protons of the D-Phe³-containing peptides outstandingly shielded at higher field.

In order to evaluate these NMR observations, Molecular Mechanics calculations were performed with the purpose of building 3D models for both opiorphin and its [D-Phe³]-analog. Due to their small size, both peptides adopted different minima by a conformational search protocol with Macromodel. However, a minimum of [D-Phe³]-opiorphin presents spatial proximity between the aromatic ring and the H γ protons of Arg², explaining the shielding at higher field observed in the D-Phe³-containing peptides. Besides, the observed NOE contacts (NH-[D-Phe³]/NH-Ser⁴ and NH-Arg⁵/H β_s -Ser⁴), as well as the vicinal coupling constant $^3J_{\text{H}\alpha-\text{H}\beta_2}$ and $^3J_{\text{H}\alpha-\text{H}\beta_3}$ for the aromatic residue D-Phe³ (4.6 Hz and 11.3 Hz respectively) are in agreement with a value for the dihedral angle χ_1 (that defines the orientation of the side chain with respect to the peptidic backbone) of around -60°. In contrast, in the native opiorphin (the L-Phe³-containing peptide), the side chain of the Phe³ residue cannot be matched with a single conformation around the dihedral angle χ_1 of Phe³. In the global minimum of opiorphin the H γ protons of Arg² are not close to the aromatic ring of Phe³. In addition, the NOE pattern recorded for opiorphin indicates a different spatial orientation of its residues. The most remarkable differences between both peptides are observed for the contacts NH-Arg⁵/H β_s -Ser⁴ and NH-Phe³/NH-Ser⁴.

All these experimental results highlight the importance of this aromatic residue in opiorphin binding activity and its specificity towards these metalloenzymes. Fluorination and perfluorination at Phe³ did not improve dual inhibition activity *versus* hNEP or hAPN, showing that these modifications at Phe³ don't induce new stabilizing interactions with the targets. However, the substitution of L-Phe³ by D-Phe³ has a

profound effect, enhancing the hAPN inhibition potency while reducing its hNEP inhibition activity. Conformational studies in solution demonstrated that the D-Phe³-containing peptide presents a cation/ π interaction between the aromatic ring of D-Phe³ and the guanidinium group of Arg², which is not present in natural opiorphin. In this study a higher-affinity selective APN inhibitor has been identified, whose potency enhancement depends on its conformational rearrangement in solution.

Chapter 2 – Structural analysis of glycopeptides of nociceptin by NMR

Nociceptin is a neuropeptide which is related to opioids. However, nociceptin doesn't interact with the classical opioid receptors¹⁸⁷. This peptide is recognized by the N/OFQ receptor (NOP), which belongs to the G-protein-coupled receptor superfamily. Their distribution is mainly in the central nervous system and in the spinal cord, but they were identified in other physiological processes in distinct systems¹⁸⁸. As its name indicates, early studies showed that nociceptin is involved in the modulation of nociception¹⁹¹. Depending on the route of administration this molecule modulates opposite effects, because antianalgesic effect was reported after intracerebroventricular administration and the analgesia of morphine is potentiated after intrathecal administration of nociceptin¹⁹⁶.

The recognition of nociceptin by NOP produces a number of intracellular processes such as adenylyl cyclase inhibition, blockage of Ca²⁺ channels, and activation of kinases and K⁺ channels¹⁹⁰⁻¹⁹¹. This interaction is very selective¹⁹⁴, since other opioids show low affinity for this receptor, while nociceptin practically doesn't bind to opioid receptors¹⁹⁵. Although nociceptin is involved in a wide range of biological functions, NOP represents a viable drug target due to this selective interaction³⁰⁰. In this chapter 2, it was demonstrated that distinct glycosylations at different parts of nociceptin interfere in its biological activity, and these effects are related to changes in the native conformation of the peptide.

Three different glycosylated analogs of nociceptin together with the parent peptide were characterized by NMR structural studies. Nociceptin, whose sequence is FGGFTGARKSARKLANQ, was glycosylated in two different sites, Thr⁵ and Ser¹⁰. Thr⁵ was glycosylated with α -D-GalNAc ([Thr⁵-O- α -D-GalNAc]-nociceptin), while Ser¹⁰ was glycosylated both with the alpha and beta anomers of GalNAc ([Thr⁵-O- α -D-GalNAc]-nociceptin and [Ser¹⁰-O- β -D-GlcNAc]-nociceptin). The combination of their biological activity with structural NMR studies allows explaining the influence of the position of these glycosylations and the combined effect of the nature of the sugar and the configuration of the glycosidic bond.

The biological activity of nociceptin and of its three glycosylated analogs was calculated by a competition radioligand binding assay with a titrated version of the parent peptide. The results showed that the most potent peptide is [Ser¹⁰-O- α -D-GalNAc]-nociceptin ($K_I = 3.81 \pm 0.19$ nM), followed by native nociceptin ($K_I =$

7.99±1.02 nM) and [Ser¹⁰-O-β-D-GlcNAc]-nociceptin ($K_I = 10.44 \pm 2.56$ nM), and being the glycosylated peptide at Thr⁵, [Thr⁵-O-α-D-GalNAc]-nociceptin, the least active one ($K_I = 54.95 \pm 9.76$ nM).

Prior to structural NMR studies, the peptides were analyzed by Far-UV circular dichroism (CD) both in water and in TFE solution. The CD spectra of these four peptides in water showed a negative Cotton effect at 200 nm, which is indicative of a random structure. On the contrary, the CD spectra in TFE solution changed to a pattern which is characteristic of α-helix, a positive Cotton effect at 195nm and two negative Cotton effects at 205nm and at 222nm. No differences were found in the CD spectra among the four species, neither in water nor in TFE solution, and thus the difference in their activity cannot be explained by this technique.

NMR studies were performed in two distinct experimental conditions in phosphate buffered solution, in the absence and in the presence of SDS micelles. SDS micelles mimic a membrane-like environment, and are commonly used in NMR as a model for the study of membrane bound biomolecules.

The NMR experiments in aqueous solutions confirmed the CD analysis, since none of these peptides showed well-defined structures. Their secondary structure appears to be random, with rapid conformational interconversion on the NMR time scale. Apart from sequential NOE contacts, very few long range NOEs were observed, just some inter-residual contacts at the C-terminus. Besides, no major changes in the chemical shifts were identified among the four peptides, other than those expected by the glycosylation. These experimental observations confirm that nociceptin and its analogs adopt disordered conformations in water.

In contrast, NMR analyses in the presence of SDS micelles show a significant increase in the number of inter-residual NOE contacts. In particular, some of these NOEs are representative of a typical α-helix pattern, and they were observed for nociceptin as well as for all the three glycopeptides. Some examples of these NOE contacts are Hβ2-Asn¹⁶/Hδ-Lys¹³, Hβ3-Asn¹⁶/Hδ-Lys¹³, Hβ3-Asn¹⁶/Hα-Lys¹³, Hα-Leu¹⁴/Hβ-Gln¹⁷, Hα-Leu¹⁴/Hγ-Gln¹⁷, Hδ-Leu¹⁴/Hγ-Gln¹⁷, Hγ-Arg¹²/Hα-Lys⁹ and Hγ-Lys⁹/Hδ-Arg¹². Besides, the ³J_{NH-Hα} coupling constants changed considerably in the presence of SDS micelles, being their values smaller than in the absence of micelles, as expected for a structure with a higher tendency to α-helix. These findings allowed us to conclude that these peptides have the tendency to form α-helix structures in SDS micelles. Thus, in a next step, NOESY spectra were employed in order to obtain the 3D structure. Peak lists from NOESY experiments, which were peak picked and integrated with the program CARA, were utilized by the CYANA program in order to generate an ensemble of conformers in agreement with the experimental results.

Following the approach described above, the 3D structure of the four peptides was obtained and compared. The least potent peptide, [Thr⁵-O-α-D-GalNAc]-nociceptin showed clear structural differences with respect to native nociceptin as a consequence of

its glycosylation at Thr⁵, that affects its conformation, as confirmed by the unambiguous distal NOE contacts observed (NH-Thr⁵/Hε21-Gln¹⁷, Hα-Thr⁵/Hε21-Gln¹⁷, Hβ-Thr⁵/Hε21-Gln¹⁷, Hγ-Thr⁵/Hε21-Gln¹⁷, NH-Ala⁷/Hε21-Gln¹⁷ and Hβ2-Phe⁴/Hε-Gln¹⁷). Despite these bent conformations, this molecule still preserves the NOE pattern of α-helix at its C-terminus, showing that this region keeps the tendency to form a helical structure from positions 14 to 17.

In the case of nociceptin and the [Ser¹⁰-O-α-D-GalNAc]-nociceptin glycopeptide, the NMR/CYANA approach confirmed their tendency to adopt α-helix from residues Ala⁷ to Gln¹⁷. The α-GalNAc glycosylation at position 10 does not disrupt the helical structure of nociceptin, but instead this glycopeptide shows an even higher α-helix tendency than the parent peptide. On the contrary, the analysis of other glycosylated peptide at position 10, [Ser¹⁰-O-β-D-GlcNAc]-nociceptin, reveals that the flexibility of this molecule is considerably higher than that of nociceptin. Although the theoretical calculations show a certain helical tendency at the C-terminus, Ser¹⁰-O-β-D-GlcNAc glycosylation evidently induces higher flexibility along the peptide chain.

In this chapter it was demonstrated that O-glycosylation of nociceptin modulates the binding affinity to its receptor, and this modulation is correlated with structural changes produced by the inclusion of the sugar moiety. NMR studies in presence of SDS micelles show that nociceptin and both glycopeptides glycosylated at position 10 adopt certain degree of helical structure, being the beta-glycosylated analog more flexible than nociceptin and the alpha glycosylated analog. The experimental NMR results for these three peptides are related to their binding affinity, because a higher α-helix tendency and less flexibility correlate with an enhancement binding affinity. Finally, the least potent glycopeptide, [Thr⁵-O-α-D-GalNAc]-nociceptin, adopts bent conformations due to its glycosylation at position 5, affecting significantly its interaction with its receptor.

Chapter 3 – Difluoroacetamide moiety as probe to detect glycan-protein interactions by NMR

Glycan-protein interactions are involved in essential events for health and disease. Even though NMR is a powerful technique, the study of complex systems is sometimes hampered by different experimental drawbacks (fast relaxation, severe ¹H signal overlapping, etc). The development of new protocols and technologies is important in order to gain insights into these molecular recognition processes. The design of new glycan-based protocols involves the synthesis and validation of novel molecules that can be used as probes. In NMR, the use of isotopes other than ¹H shows several benefits from a technical point of view, so the application of fluorinated molecules as NMR probes has been successfully exploited in recent reported examples. The proposal of this chapter is the application of a difluoroacetamide moiety as a sensor for monitoring the interaction between GlcNAc-containing molecules and their targets. The difluoroacetamide group is easy to introduce in these glycans and it enables a robust and dual detection by both ¹H-NMR and ¹⁹F-NMR.

The evaluation of this new probe by NMR was achieved by using two difluoroacetamide-containing analogues of *N,N'*-diacetyl chitobiose and *N,N',N''*-triacetyl chitotriose. These molecules were tested *versus* wheat germ agglutinin (WGA), which is one of the most studied carbohydrate binding proteins. This model protein contains four hevein-like domains and it preferentially binds to GlcNAc moieties through hydrogen bonds and CH- π interactions.

Firstly, both molecules were fully characterized in the free state by ^1H -NMR and ^{19}F -NMR. Although the signals in the ^1H NMR spectra are highly overlapped, the spectra could be fully assigned. The protons at the difluoroacetyl moieties resonate at an isolated spectral region as well defined triplets. On the other hand, the interpretation of the ^{19}F NMR spectra is much simpler. Surprisingly, the signals of two geminal fluorine atoms at each residue are magnetically different, despite its location far away from the chiral sugar ring. As the chemical shift difference between the two coupled fluorine signals is much smaller than their strong coupling constants ($^2J_{\text{F,F}} \sim 300\text{Hz}$), these doublets are affected by a pronounced 'roof effect'. The assignment of the ^{19}F signals for each GlcNAc moiety was achieved by ^{19}F -COSY and ^1H - ^{19}F heteronuclear experiments.

Afterwards, the interaction of both probes with WGA was examined by ^1H -NMR and ^{19}F -NMR experiments. The selective line broadening observed both in ^1H and in ^{19}F in the corresponding spectra for the fluorinated *N,N'*-diacetyl chitobiose clearly demonstrated the existence of a molecular recognition process in which there is a preferred interaction with the residue at the non-reducing end. ^1H -STD and ^{19}F -STD experiments corroborated these results. Thus, diverse NMR strategies permitted to identify and characterize with epitope information the interaction of this chitobiose probe with WGA. The same procedure was applied to monitor the interaction of the fluorinated *N,N',N''*-triacetyl chitotriose with WGA. Again, selective line broadening was observed, being more pronounced for the central and non-reducing end residues, than for the reducing end one, indicating a preferred recognition process. Even though ^1H and ^{19}F NMR spectra clearly identified this chitotriose probe as a binder, STD experiments were not positive. Still, ^{19}F -DOSY experiments showed a drastic increase in the diffusion coefficient of the trisaccharide in the presence of WGA, indicating that this ^{19}F -tagged chitotriose is recognized with higher affinity than the corresponding ^{19}F -tagged chitobiose, as it was reported also for non-fluorinated parent saccharides.

A molecular modeling protocol was applied in order to understand the intermolecular forces which are involved in the molecular recognition of the CF_2HCO moiety. First, simple models for the complex between four different acetamide groups, which possess the formula $\text{CF}_x\text{H}_y\text{-CO-NH-CH}_3$, with benzene and *p*-methyl phenol (as simplified models of phenylalanine and tyrosine respectively) were built and evaluated by DFT methods (Gaussian 09). The initial structures for these calculations were prepared by placing the $\text{CF}_x\text{H}_y\text{-}$ group stacked on top of the corresponding aromatic ring providing either CH/ π or CF/ π interactions. The results showed the same trends for both

simplified models of the aromatic residues. The interaction energy of the complex depends on the degree of fluorination of the acetamide group (the values X and Y in CF_XH_Y). The ones containing CF_2H and CFH_2 groups showed the lowest interaction energy when it is the CH bond the one oriented towards the aromatic ring, in contrast to the situation in which it is the CF bond the one stacked on top of the aromatic ring. In the case of the CF_2H group, the CH/π interaction is 1 kcal/mol more favorable than the same interaction for the regular methyl group, and significantly better than the CF/π interaction for CF_3 group (2.5 kcal/mol). These energy values are in agreement with the geometries obtained for these complexes, since better interaction energy is translated into shorter distances between the acetamide and the centroid of the aromatic ring. These results show that the CF_2H -containing acetamide moiety mimics the natural methyl group and that its CH/π interactions are slightly more favorable than that for the natural acetamide. Finally, both difluoroacetamide-containing analogues of N,N' -diacetyl chitobiose and N,N',N'' -triacetyl chitotriose were docked into one of the WGA binding sites, and the 3D structural models were submitted to short molecular dynamics (MD) runs. The molecular recognition features of both complexes are the same as for the parent di and trisaccharides, which are well established: in the case of chitobiose, the non-reducing end residue is positioned at hevein subsite +1, establishing interactions with residues Ser¹⁹, Trp²³ and Tyr³⁰, while the reducing end residue is at subsite +2, interacting with Trp²¹. In contrast, chitotriose is recognized in two different manners. The binding mode A places the non-reducing residue at subsite +1 and the adjacent central residue at subsite +2 (as the chitobiose pose described above), while the binding mode B locates the central sugar at subsite +1, and the terminal reducing end residue is at subsite +2. All the modeled complexes with the difluoroacetamide-containing molecules were stable along the MD simulation and their interactions are analogous to those described above. These computational data are in agreement with our NMR observations described above. Fluorinated chitobiose establishes major contacts by its non-reducing moiety and chitotriose by the non-reducing and central residues.

In chapter 3 it was demonstrated that the introduction of the difluoroacetamide group is a practical tagging method for analyzing the interaction of NAc-containing glycans with receptors by NMR. This tag can be employed as an NMR probe, and its dual detection by ¹H-NMR and ¹⁹F-NMR allows monitoring the molecular recognition processes through a higher diversity of NMR experiments. Besides, these studies showed that the difluoroacetamide tag is effective throughout a wide range of binding affinities, and behaves similarly to the natural acetamide fragment. This strategy is a powerful tool which can be utilized in interacting systems involving *N*-Acetylated monosaccharides, such as GlcNAc, GalNAc or NeuNAc, that are found widespread in natural glycans.

Chapter 4 – *E. coli* β-galactosidase inhibitors through modifications at the aglyconic moiety

Glycosidases control diverse processes in critical biological events, and thus understanding the mechanism of action of their inhibitors provides essential information

for a structure-based drug design. The catalytic cleavage of a glycosidic linkage is preceded by a molecular distortion of the substrate upon binding to the enzyme. In fact, numerous enzymes recognize their substrate in high-energy conformations. These conformers, which are minor populated in solution, possess high-energy shapes of the six-membered ring at their non-reducing moiety or adopt high-energy geometries around their glycosidic angles^{97,290}.

E. coli β -galactosidase is considered as a model protein of glycosidase²⁷⁸. The key aspects of its mechanism have been deeply investigated by X-ray crystallography and site-directed mutagenesis, and its catalytic and recognition residues have been characterized in detailed³⁰¹. Prior to catalysis, the enzyme induces either a significant deformation of the pyranose ring from a 4C_1 chair to a half-chair or sofa conformations, or a distortion around the glycosidic torsions, rotating ϕ from a *syn* to an *anti* conformation¹¹⁷. The reported studies so far show that the changes take place at the non-reducing end residue, and the adopted shape place the aglycone moiety outside the catalytic site. In this chapter however, we show, through the employment of two new inhibitors, that the affinity of the β -galactosidase inhibitors can be modulated through modifications at their aglyconic moiety.

The two new inhibitors are thiodisaccharides: inhibitor 1 is *benzyl 3-deoxy-4-S-(β -D-galactopyranosyl)-4-thio- β -D-erythro-pentopyranoside* and inhibitor 2 is *benzyl 3-deoxy-4-S-(β -D-galactopyranosyl)-4-thio- β -D-threo-pentopyranoside*. Their interaction with *E. coli* β -galactosidase was analyzed by NMR ligand-based methods in order to rationalize the difference in their affinity. Although these thiodisaccharides differ only in the stereochemistry at position C2 of the pentopyranose (the aglycone) at the reducing-end, the difference in their inhibition constants was reported to be 25-fold (Inhibitor 1 K_I =800 μ M, and Inhibitor 2 K_I =32 μ M).

Conformational analysis of these molecules in buffer solution was first carried out. The standard *J*-coupling constant analysis and 2D NOESY experiments showed conformational differences between the two inhibitors. The galactopyranose residue of both molecules, the non-reducing moiety, adopts the typical 4C_1 chair as expected. However, the pentopyranose rings show a conformational equilibrium between two chair conformations, 4C_1 and 1C_4 . The *J*-coupling constant analysis allows to establish that for inhibitor 1, the aglycon is mainly in a 4C_1 conformation (4C_1 , 91%; 1C_4 , 9%), while for inhibitor 2 there is a 1:1 conformational equilibrium (4C_1 , 47%; 1C_4 , 53%). The preference for a 4C_1 chair in the aglycon of 1 was confirmed by 2D NOESY experiments. On the other hand, for inhibitor 2 intrarresidual NOEs characteristic of both chairs were observed in the NOESY spectrum: the 4C_1 chair was identified by the NOEs H1-H3B and H3B-H5B, while the 1C_4 chair by the NOE H3A-H5A. In order to determine the conformation around the thioglycosidic linkage a combined molecular mechanics/NMR approach was applied. Theoretical calculations suggested the existence of three possible energy minima when the pentopyranose adopts the 4C_1 conformation and two other minima when it is in a 1C_4 conformation. By inspection of the

interresidual NOE contacts, it was concluded that the population distribution for inhibitor 1 contains two or three rotamers, always with the aglycone in the 4C_1 conformation. For inhibitor 2, the conformational ensemble is composed of three to five conformers, where the aglycone adopts both possible chair geometries. Due to signal overlapping some conformers couldn't be unambiguously confirmed or excluded.

The molecular recognition of both inhibitors by *E.coli* β -galactosidase was confirmed by STD and TR-NOESY experiments. Apart from validating the enzyme-ligand interaction, these NMR experiments provided information about the binding epitope and the conformational geometries adopted in the bound state. STD analysis suggested that 1 and 2 display distinct binding modes. For compound 1 the largest STD percentages correspond to the galactose H4' and aromatic protons, but for 2 the major transfer occurred for the aromatic protons. In compound 2, the protons of the Galactose residue, especially H2', also receives significant magnetization transfer, whereas the protons of the pentopyranose show less STD intensity. In contrast, for molecule 1 the aglyconic residue show important STD effect, much more than H2' and H3' at non-reducing end. The STD pattern of the galactose residues of these inhibitors was compared with that of IPTG, a known β -galactosidase binder, showing that it is more similar to that of inhibitor 2.

TR-NOESY experiments also corroborated the binding of both inhibitors: the NOE cross peaks for both ligands in the presence of the enzyme show the same sign as diagonal peaks, indicating that they are negative NOE (and thus transferred NOE, trNOE). Key differences were observed in the pattern of the cross peaks between the free and bound states for both inhibitors, revealing that the bound conformations differ from the population distributions in solution. For instance, inhibitor 1 adopts both 4C_1 and 1C_4 chairs at its pentopyranose ring, contrary to situation found for the free state. Besides, the trNOE analysis permitted to determine that the bound conformer with a 4C_1 chair has an *anti* ϕ /*syn* ψ interglycosidic conformation, while the conformer with 1C_4 chair shows an *syn* ϕ /*syn* ψ interglycosidic conformation. Thus, the ground state conformation found for 1 in the free state is not recognized by the enzyme, but two alternative high-energy conformations. A similar analysis for the interaction with inhibitor 2 also suggested two different binding conformers, which coincide with the two bound conformations found for 1. However, a distinct conformational selection process occurs for the recognition of these two inhibitors. Contrary to the case of inhibitor 1, for inhibitor 2, the two recognized conformers already existed in the conformational ensemble in the free state. Thus, for both inhibitors two conformations are bound: the *anti* ϕ /*syn* ψ with a 4C_1 chair for the aglyconic moiety and the *syn* ϕ /*syn* ψ with a 1C_4 chair. Therefore, β -galactosidase distorts the mayor conformations in the free state, *syn* ϕ /*syn* ψ with a 4C_1 chair at the aglycone, either by turning from *syn* ϕ to *anti* ϕ at the glycosidic bond, or by changing from a 4C_1 to a 1C_4 chair at the aglycone moiety.

Molecular docking and Molecular Dynamics (MD) simulations were performed in order to provide plausible three-dimensional structures for the complexes which could explain the NMR experimental data. Docking studies were performed with the software Glide, without any constraint, and using the pdb structure 1JYY for the *E.coli* β -galactosidase structure. Both inhibitors, 1 and 2, were docked in two distinct starting geometries, with the pentopyranose ring in 4C_1 and in 1C_4 chair conformations. Due to the lack of constraints, transitions between the two chair conformations of the aglyconic moiety, as well as rotation around the thioglycosidic linkage were allowed in the calculation. The resulting 3D population histogram for the docking of molecule 2 reveals that the most populated conformer is the *anti* ϕ /*syn* ψ with 4C_1 chair at the aglycone, and the second most populated is the *syn* ϕ /*syn* ψ with 1C_4 chair. These predictions are in full agreement with the experimental NMR data.

Docking analysis considers the protein as a rigid entity. Thus, in order to test the stability of the obtained complexes in the context of protein flexibility, unrestrained MD simulations were performed. Along 10ns MD simulations, the protein structures were stable and the ligands didn't diffuse into the solvent. The glycosidic torsion angles were fairly stable. For ϕ , the initial orientation was maintained all throughout the MD trajectories. For ψ , although certain fluctuations were observed, its value was always kept within the *syn* ψ region. The intermolecular interactions taken place for each of the complexes were analyzed in detail and compared. In fact, substantial differences were identified.

CH- π interactions, a relevant feature in substrate recognition by glycosidases²⁸¹, were monitored by measuring the distances between the CH pyranose protons and the centroids of the aromatic ring of key tryptophan residues (Trp⁵⁶⁸ and Trp⁹⁹⁹). For instance, the distance between the Galactose protons and Trp⁵⁶⁸ centroid is practically constant for all simulation. Therefore, this essential carbohydrate-aromatic stacking interaction is maintained along the MD simulations. The complexes are further stabilized by several hydrogen bonds, that are known to attach the galactosyl non-reducing unit into the active site (GalO2/Glu⁴⁶¹, GalO3/Glu⁵³⁷, GalO4/Asp²⁰¹, GalO6/His⁵⁶⁰, GalO6/Asn⁶⁰⁴, GalO6/Asp²⁰¹).

As the galactose residue is fixed, the glycosidic torsion angles mainly determine the position of the aglycone moiety in the complexes. Thereby, in the *anti* ϕ /*syn* ψ with 4C_1 chair conformation, the pentopyranose ring resembles the orientation of the departing leaving group at the end of the enzymatic process, while in the *syn* ϕ /*syn* ψ geometry with the alternative 1C_4 chair conformation; the aglycon is adequately accommodated without major perturbation of the binding site. In the case of the *syn* ϕ /*syn* ψ geometry, for both inhibitors the three protons H3A, H4 and H5A of the aglycon moiety are stacked on top of Trp⁹⁹⁹. No such interaction can be established in the case of the *anti* ϕ /*syn* ψ geometry. The position of the aromatic ring attached to the reducing end was also different for the two recognized conformers. In the case of the *syn* ϕ /*syn* ψ 1C_4 conformer the phenyl ring interacts temporarily with Phe⁵¹² by aromatic-aromatic

interactions, while in the $\text{anti}\phi/\text{syn}\psi$ 4C_1 geometry, it establishes transient cation- π interactions with Arg⁸⁰⁰. In general, for inhibitor 1 both recognized conformers show smaller number of interactions at pentopyranose ring than in binder 2, so these *in-silico* models agree with the weaker affinity of 1 with respect to 2.

As mentioned above, the STD experiments indicate that the binding modes of molecule 1 and 2 should be different. In fact, inhibitor 1 has got a non-competitive inhibition profile, while inhibitor 2 show a stronger pure competitive inhibition profile, so all these structural discrepancies between 1 and 2 could explain these distinct activity profiles.

In this chapter, the conformation analysis of two inhibitors against *E. coli* β -galactosidase explains the difference between their inhibitory activities. Although between these molecules there is a single change in the stereochemistry at one stereogenic center at aglyconic moiety, there is a 25-fold difference in the inhibition constant. NMR studies showed that, in the free state, their conformation behavior is quite different. However, in the bound state, the same two conformers are recognized for both inhibitors. One of this conformers has a $\text{syn}\phi/\text{syn}\psi$ geometry around thioglycosidic linkage and a 1C_4 chair shape at the aglycone, while the other has a $\text{anti}\phi/\text{syn}\psi$ conformation with 4C_1 chair at the aglycone. When bound to the enzyme, the most potent inhibitor, 2, adopts mainly the $\text{syn}\phi/\text{syn}\psi$ geometry with 1C_4 conformation, while molecule 1 adopts both conformations in a similar ratio. None of the inhibitors behave as *O*-glycoside lactose analogs, because the β -galactosidase selects two unusual conformations, either around the ϕ glycosidic angle or at the aglyconic moiety. Thus, a distorted geometry at aglyconic moiety in free state determines an enhancement of inhibition activity. The differences in the inhibition of these molecules show the interest of a new generation of molecules which can modulate their activity by modifying their aglyconic moiety at selected positions

Bibliografía

- (1) Wüthrich, K. *NMR of Proteins and Nucleic Acids* **1986**.
- (2) Roberts, G. C. K. *NMR of Macromolecules: A Practical Approach* **1993**.
- (3) Stassinopoulou, C. I. *NMR of Biological Macromolecules* **1994**.
- (4) Evans, J. N. S. *Biomolecular NMR Spectroscopy* **1995**.
- (5) Güntert, P. *Quarterly Reviews of Biophysics* **1998**, 31, 145.
- (6) Jiménez-Barbero, J.; Peters, T. *NMR Spectroscopy of Glycoconjugates* **2003**.
- (7) Zerbe, O.; Mannhold, R.; Kubinyi, H.; Folkers, G. *BioNMR in Drug Research* **2003**.
- (8) Zartler, E.; Shapiro, M. *Fragment-based Drug Discovery: A Practical Approach* **2008**.
- (9) Dunbrack, R. L.; Karplus, M. *Nature Structural & Molecular Biology* **1994**, 1, 334.
- (10) Zhou, A. Q.; O'Hern, C. S.; Regan, L. *Proteins: Structure, Function, and Bioinformatics* **2014**, 82, 2574.
- (11) Gómez-Moreno Calera, C.; Sancho Sanz, J. In *Estructura de Proteínas*; Ed. Arial: **2003**.
- (12) Chang, X.; Keller, D.; Bjørn, S.; Led, J. J. *Magnetic Resonance in Chemistry* **2001**, 39, 477.
- (13) Bauer, F.; Schweimer, K.; Klüver, E.; Conejo-Garcia, J.-R.; Forssmann, W.-G.; Rösch, P.; Adermann, K.; Sticht, H. *Protein Science* **2001**, 10, 2470.
- (14) Conibear, A. C.; Rosengren, K. J.; Harvey, P. J.; Craik, D. J. *Biochemistry* **2012**, 51, 9718.
- (15) <http://www.nature.com/scitable/definition/peptide-317>.
- (16) Sanger, F. *Biochemical Journal* **1945**, 39, 507.
- (17) Sanger, F. *Biochemical Journal* **1949**, 45, 563.
- (18) Sanger, F.; Tuppy, H. *Biochemical Journal* **1951**, 49, 463.
- (19) Sanger, F.; Tuppy, H. *Biochemical Journal* **1951**, 49, 481.
- (20) Sanger, F.; Thompson, E. O. P. *Biochemical Journal* **1953**, 53, 353.
- (21) Sanger, F.; Thompson, E. O. P. *Biochemical Journal* **1953**, 53, 366.
- (22) Adams, M. J.; Blundell, T. L.; Dodson, E. J.; Dodson, G. G.; Vijayan, M.; Baker, E. N.; Harding, M. M.; Hodgkin, D. C.; Rimmer, B.; Sheat, S. *Nature* **1969**, 224, 491.
- (23) Schaffer, L. *European Journal of Biochemistry* **1994**, 221, 1127.
- (24) Yip, C. C.; Ottensmeyer, P. *Journal of Biological Chemistry* **2003**, 278, 27329.
- (25) Sasaki, K.; Dockerill, S.; Adamiak, D. A.; Tickle, I. J.; Blundell, T. *Nature* **1975**, 257, 751.
- (26) Braun, W.; Wider, G.; Lee, K. H.; Wüthrich, K. *Journal of Molecular Biology* **1983**, 169, 921.
- (27) Siu, F. Y.; He, M.; de Graaf, C.; Han, G. W.; Yang, D.; Zhang, Z.; Zhou, C.; Xu, Q.; Wacker, D.; Joseph, J. S.; Liu, W.; Lau, J.; Cherezov, V.; Katritch, V.; Wang, M.-W.; Stevens, R. C. *Nature* **2013**, 499, 444.
- (28) Cornet, B.; Bonmatin, J.-M.; Hetru, C.; Hoffmann, J. A.; Ptak, M.; Vovelle, F. *Structure*, 3, 435.
- (29) Janssen, B. J. C.; Schirra, H. J.; Lay, F. T.; Anderson, M. A.; Craik, D. J. *Biochemistry* **2003**, 42, 8214.
- (30) Rose, G. D.; Gierasch, L. M.; Smith, J. A. *Adv. Protein Chem.* **1985**, 37, 1.
- (31) Whitby, L. R.; Ando, Y.; Setola, V.; Vogt, P. K.; Roth, B. L.; Boger, D. L. *Journal of the American Chemical Society* **2011**, 133, 10184.

- (32) Eckhardt, B.; Grosse, W.; Essen, L.-O.; Geyer, A. *Proceedings of the National Academy of Sciences* **2010**, *107*, 18336.
- (33) Sands, B. E.; Podolsky, D. K. *Annual Review of Physiology* **1996**, *58*, 253.
- (34) Thim, L. *CMLS, Cell. Mol. Life Sci.* **1997**, *53*, 888.
- (35) McAuliffe, O.; Ross, R. P.; Hill, C. *Lantibiotics: structure, biosynthesis and mode of action*; FEMS Microbiology Reviews, 2001; Vol. 25.
- (36) Willey, J. M.; van der Donk, W. A. *Annual Review of Microbiology* **2007**, *61*, 477.
- (37) Altona, C.; Knobler, C.; Romers, C. *Acta Crystallographica* **1963**, *16*, 1217.
- (38) Juaristi, E.; Cuevas, G. *The anomeric effect*, Boca Raton, CRC Press. **1995**.
- (39) Lemieux, R. U.; Koto, S.; Voisin, D. In *Anomeric Effect*; American Chemical Society: 1979; Vol. 87, p 17.
- (40) Lemieux, R. U.; Koto, S. *Tetrahedron* **1974**, *30*, 1933.
- (41) Cremer, D.; Pople, J. A. *Journal of the American Chemical Society* **1975**, *97*, 1354.
- (42) Jeffrey, G. A.; Yates, J. H. *Carbohydrate Research* **1979**, *74*, 319.
- (43) Casu, B.; Petitou, M.; Provasoli, M.; Sinaÿ, P. *Trends in Biochemical Sciences* **1988**, *13*, 221.
- (44) Lee, Y. C.; Lee, R. T. *Accounts of Chemical Research* **1995**, *28*, 321.
- (45) Lee, Y. C. *The FASEB Journal* **1992**, *6*, 3193.
- (46) Fernández-Alonso, M. d. C.; Díaz, D.; Berbis, M. Á.; Marcelo, F.; Cañada, J.; Jiménez-Barbero, J. *Current Protein & Peptide Science* **2012**, *13*, 816.
- (47) Roldós, V.; Cañada, F. J.; Jiménez-Barbero, J. *ChemBioChem* **2011**, *12*, 990.
- (48) Słomińska-Wojewódzka, M.; Sandvig, K. *Molecules* **2015**, *20*, 9816.
- (49) Ley, K. *Trends in Molecular Medicine*, *9*, 263.
- (50) Coombe, D. R.; Kett, W. C. *Cellular and Molecular Life Sciences* **2005**, *62*, 410.
- (51) Elgavish, S.; Shaanan, B. *Journal of Molecular Biology* **1998**, *277*, 917.
- (52) Quiocho, F. A. *Probing the atomic interactions between proteins and carbohydrates*; Biochemical Society Transactions, 1993; Vol. 21.
- (53) Chávez, M. I.; Andreu, C.; Vidal, P.; Aboitiz, N.; Freire, F.; Groves, P.; Asensio, J. L.; Asensio, G.; Muraki, M.; Cañada, F. J.; Jiménez-Barbero, J. *Chemistry – A European Journal* **2005**, *11*, 7060.
- (54) Lemieux, R. U. *Accounts of Chemical Research* **1996**, *29*, 373.
- (55) Asensio, J. L.; Ardá, A.; Cañada, F. J.; Jiménez-Barbero, J. *Accounts of Chemical Research* **2013**, *46*, 946.
- (56) Wright, C. S. *Journal of Molecular Biology* **1990**, *215*, 635.
- (57) Asensio, J. L.; Canada, F. J.; Bruix, M.; Rodríguez-Romero, A.; Jimenez-Barbero, J. *European Journal of Biochemistry* **1995**, *230*, 621.
- (58) Asensio, J. L.; Cañada, F. J.; Siebert, H. C.; Laynez, J.; Poveda, A.; Nieto, P. M.; Soedjanaamadja, U. M.; Gabius, H. J.; Jiménez-Barbero, J. *Chemistry and Biology* **2000**, *7*, 529.
- (59) Weis, W.; Brown, J. H.; Cusack, S.; Paulson, J. C.; Skehel, J. J.; Wiley, D. C. *Nature* **1988**, *333*, 426.
- (60) Muñoz, E. M.; López de la Paz, M.; Jiménez-Barbero, J.; Ellis, G.; Pérez, M.; Vicent, C. *Chemistry – A European Journal* **2002**, *8*, 1908.
- (61) Vicente, V.; Martin, J.; Jiménez-Barbero, J.; Chiara, J. L.; Vicent, C. *Chemistry – A European Journal* **2004**, *10*, 4240.
- (62) Hodel, A. E.; Quiocho, F. A. *International Tables for Crystallography* **2006**, *F*, 579.
- (63) Muraki, M.; Ishimura, M.; Harata, K. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2002**, *1569*, 10.

- (64) Varghese, J. N.; Colman, P. M.; van Donkelaar, A.; Blick, T. J.; Sahasrabudhe, A.; McKimm-Breschkin, J. L. *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 11808.
- (65) Panigrahi, S. K.; Desiraju, G. R. *Proteins: Structure, Function, and Bioinformatics* **2007**, *67*, 128.
- (66) Quijcho, F. A. *Annual Review of Biochemistry* **1986**, *55*, 287.
- (67) Merritt, E. A.; Sarfaty, S.; Akker, F. V. D.; L'Hoir, C.; Martial, J. A.; Hol, W. G. J. *Protein Science* **1994**, *3*, 166.
- (68) Watowich, S. J.; Skehel, J. J.; Wiley, D. C. *Structure*, *2*, 719.
- (69) Stehle, T.; Yan, Y.; Benjamin, T. L.; Harrison, S. C. *Nature* **1994**, *369*, 160.
- (70) Stehle, T.; Harrison, S. C. *Structure* **1996**, *4*, 183.
- (71) Sainz-Polo, M. A.; Valenzuela, S. V.; González, B.; Pastor, F. I. J.; Sanz-Aparicio, J. *Journal of Biological Chemistry* **2014**, *289*, 31088.
- (72) Xu, X.; Zhu, X.; Dwek, R. A.; Stevens, J.; Wilson, I. A. *Journal of Virology* **2008**, *82*, 10493.
- (73) Oakley, A. J.; Barrett, S.; Peat, T. S.; Newman, J.; Streltsov, V. A.; Waddington, L.; Saito, T.; Tashiro, M.; McKimm-Breschkin, J. L. *Journal of Medicinal Chemistry* **2010**, *53*, 6421.
- (74) Escuret, V.; Collins, P. J.; Casalegno, J.-S.; Vachieri, S. G.; Cattle, N.; Ferraris, O.; Sabatier, M.; Frobert, E.; Caro, V.; Skehel, J. J.; Gamblin, S.; Valla, F.; Valette, M.; Ottmann, M.; McCauley, J. W.; Daniels, R. S.; Lina, B. *Journal of Infectious Diseases* **2014**, *210*, 1260.
- (75) Stierand, K.; Rarey, M. *ACS Medicinal Chemistry Letters* **2010**, *1*, 540.
- (76) Sixma, T. K.; Pronk, S. E.; Kalk, K. H.; van Zanten, B. A. M.; Berghuis, A. M.; Hol, W. G. J. *Nature* **1992**, *355*, 561.
- (77) Merritt, E. A.; Sixma, T. K.; Kalk, K. H.; van Zanten, B. A. M.; Hol, W. G. J. *Molecular Microbiology* **1994**, *13*, 745.
- (78) Holmner, Å.; Lebens, M.; Teneberg, S.; Ångström, J.; Ökvist, M.; Krengel, U. *Structure*, *12*, 1655.
- (79) Merritt, E. A.; Pronk, S. E.; Sixma, T. K.; Kalk, K. H.; van Zanten, B. A. M.; Hol, W. G. J. *FEBS Letters*, *337*, 88.
- (80) Gregoriou, M.; Noble, M. E. M.; Watson, K. A.; Garman, E. F.; Johnson, L. N.; Krulle, T. M.; Fuetene, C. D. L.; Fleet, G. W. J.; Oikonomakos, N. G. *Protein Science* **1998**, *7*, 915.
- (81) Archontis, G.; Watson, K. A.; Xie, Q.; Andreou, G.; Chrysina, E. D.; Zographos, S. E.; Oikonomakos, N. G.; Karplus, M. *Proteins: Structure, Function, and Bioinformatics* **2005**, *61*, 984.
- (82) Loris, R.; Hamelryck, T.; Bouckaert, J.; Wyns, L. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **1998**, *1383*, 9.
- (83) Zelensky, A. N.; Gready, J. E. *FEBS Journal* **2005**, *272*, 6179.
- (84) Weis, W. I.; Drickamer, K.; Hendrickson, W. A. *Nature* **1992**, *360*, 127.
- (85) Feinberg, H.; Park-Snyder, S.; Kolatkar, A. R.; Heise, C. T.; Taylor, M. E.; Weis, W. I. *Journal of Biological Chemistry* **2000**, *275*, 21539.
- (86) Ng, K. K.-S.; Drickamer, K.; Weis, W. I. *Journal of Biological Chemistry* **1996**, *271*, 663.
- (87) Lavie, A.; Allen, K. N.; Petsko, G. A.; Ringe, D. *Biochemistry* **1994**, *33*, 5469.
- (88) Weignerova, L.; Bojarova, P.; Kren, V. In *Carbohydrate Chemistry: Volume 35*; The Royal Society of Chemistry: 2009; Vol. 35, p 311.
- (89) <http://www.cazy.org/Glycoside-Hydrolases.html>.
- (90) Henrissat, B. *Biochemical Journal* **1991**, *280*, 309.
- (91) Henrissat, B.; Bairoch, A. *Biochemical Journal* **1993**, *293*, 781.
- (92) Henrissat, B.; Bairoch, A. *Biochemical Journal* **1996**, *316*, 695.
- (93) Vasella, A.; Davies, G. J.; Böhm, M. *Current Opinion in Chemical Biology* **2002**, *6*, 619.

- (94) Davies, G. J.; Gloster, T. M.; Henrissat, B. *Current Opinion in Structural Biology* **2005**, *15*, 637.
- (95) Mark, B. L.; Vocadlo, D. J.; Knapp, S.; Triggs-Raine, B. L.; Withers, S. G.; James, M. N. G. *Journal of Biological Chemistry* **2001**, *276*, 10330.
- (96) Yip, V. L. Y.; Withers, S. G. *Current Opinion in Chemical Biology* **2006**, *10*, 147.
- (97) Rye, C. S.; Withers, S. G. *Current Opinion in Chemical Biology* **2000**, *4*, 573.
- (98) Vuong, T. V.; Wilson, D. B. *Biotechnology and Bioengineering* **2010**, *107*, 195.
- (99) Rempel, B. P.; Withers, S. G. *Glycobiology* **2008**, *18*, 570.
- (100) Withers, S. G.; Street, I. P.; Bird, P.; Dolphin, D. H. *Journal of the American Chemical Society* **1987**, *109*, 7530.
- (101) Withers, S. G.; Rupitz, K.; Street, I. P. *Journal of Biological Chemistry* **1988**, *263*, 7929.
- (102) Ferrer, M.; Golyshina, Olga V.; Plou, Francisco J.; Timmis, Kenneth N.; Golyshin, Peter N. *Biochemical Journal* **2005**, *391*, 269.
- (103) Watts, A. G.; Withers, S. G. *Canadian Journal of Chemistry* **2004**, *82*, 1581.
- (104) Tull, D.; Burgoyne, D. L.; Chow, D. T.; Withers, S. G.; Aebersold, R. *Analytical Biochemistry* **1996**, *234*, 119.
- (105) Chir, J.; Withers, S.; Wan, C.-F.; Li, Y.-K. *Biochemical Journal* **2002**, *365*, 857.
- (106) Vocadlo, D. J.; Wicki, J.; Rupitz, K.; Withers, S. G. *Biochemistry* **2002**, *41*, 9736.
- (107) Daniels, L. B.; Glew, R. H.; Radin, N. S.; Vunnam, R. R. *Clinica Chimica Acta* **1980**, *106*, 155.
- (108) Zhao, H.; Grabowski, G. A. *Cellular and Molecular Life Sciences* **2002**, *59*, 694.
- (109) Butters, T. D. *Current Opinion in Chemical Biology* **2007**, *11*, 412.
- (110) Wolfenden, R. *Annual Review of Biophysics and Bioengineering* **1976**, *5*, 271.
- (111) Gloster, T. M.; Davies, G. J. *Organic & Biomolecular Chemistry* **2010**, *8*, 305.
- (112) Wolfenden, R.; Snider, M.; Ridgway, C.; Miller, B. *Journal of the American Chemical Society* **1999**, *121*, 7419.
- (113) Snider, M. J.; Gaunitz, S.; Ridgway, C.; Short, S. A.; Wolfenden, R. *Biochemistry* **2000**, *39*, 9746.
- (114) Gloster, T. M.; Meloncelli, P.; Stick, R. V.; Zechel, D.; Vasella, A.; Davies, G. J. *Journal of the American Chemical Society* **2007**, *129*, 2345.
- (115) Lillielund, V. H.; Jensen, H. H.; Liang, X.; Bols, M. *Chemical Reviews* **2002**, *102*, 515.
- (116) Espinosa, J. F.; Montero, E.; Vian, A.; García, J. L.; Dietrich, H.; Schmidt, R. R.; Martín-Lomas, M.; Imberty, A.; Cañada, F. J.; Jiménez-Barbero, J. *Journal of the American Chemical Society* **1998**, *120*, 1309.
- (117) García-Herrero, A.; Montero, E.; Muñoz, J. L.; Espinosa, J. F.; Vián, A.; García, J. L.; Asensio, J. L.; Cañada, F. J.; Jiménez-Barbero, J. *Journal of the American Chemical Society* **2002**, *124* 4804.
- (118) Montero, E.; Vallmitjana, M.; Pérez-Pons, J. A.; Querol, E.; Jiménez-Barbero, J.; Cañada, F. J. *FEBS Letters* **1998**, *421*, 243.
- (119) Wüthrich, K.; Nagayama, K. *European Journal of Biochemistry* **1981**, *115*, 653.
- (120) Krishnan, V. V.; Rupp, B. In *eLS*; John Wiley & Sons, Ltd: 2001.
- (121) Garbuzynskiy, S. O.; Melnik, B. S.; Lobanov, M. Y.; Finkelstein, A. V.; Galzitskaya, O. V. *Proteins: Structure, Function, and Bioinformatics* **2005**, *60*, 139.
- (122) Claridge, T. D. W. In *Tetrahedron Organic Chemistry Series*; Timothy, D. W. C., Ed.; Elsevier: 2009; Vol. Volume 27, p 11.
- (123) Jacobsen, N. E. In *NMR Spectroscopy Explained*; John Wiley & Sons, Inc.: 2007, p 1.
- (124) Haasnoot, C. A. G.; de Leeuw, F. A. A. M.; Altona, C. *Tetrahedron* **1980**, *36*, 2783.

- (125) Lemieux, R. U.; Kullnig, R. K.; Bernstein, H. J.; Schneider, W. G. *Journal of the American Chemical Society* **1958**, *80*, 6098.
- (126) Claridge, T. D. W. In *Tetrahedron Organic Chemistry Series*; Timothy, D. W. C., Ed.; Elsevier: 2009; Vol. Volume 27, p 129.
- (127) Jacobsen, N. E. In *NMR Spectroscopy Explained*; John Wiley & Sons, Inc.: 2007, p 353.
- (128) Claridge, T. D. W. In *Tetrahedron Organic Chemistry Series*; Timothy, D. W. C., Ed.; Elsevier: 2009; Vol. Volume 27, p 247.
- (129) Jacobsen, N. E. In *NMR Spectroscopy Explained*; John Wiley & Sons, Inc.: 2007, p 408.
- (130) Wüthrich, K. *Biopolymers* **1983**, *22*, 131.
- (131) Güntert, P. In *Protein NMR Techniques*; Downing, A. K., Ed.; Humana Press: 2004; Vol. 278, p 353.
- (132) Keller, R. L. J. In *Optimizing the process of nuclear magnetic resonance spectrum analysis and computer aided resonance assignment*; ETH: 2005.
- (133) Poveda, A.; Jiménez-Barbero, J. *Chemical Reviews* **1998**, *27*, 133.
- (134) Viegas, A.; Macedo, A.; Cabrita, E. In *Ligand-Macromolecular Interactions in Drug Discovery*; Roque, A. C. A., Ed.; Humana Press: 2010; Vol. 572, p 81.
- (135) Unione, L.; Galante, S.; Diaz, D.; Canada, F. J.; Jimenez-Barbero, J. *MedChemComm* **2014**, *5*, 1280.
- (136) Mayer, M.; Meyer, B. *Angewandte Chemie. International Ed. In English* **1999**, *38*, 1784.
- (137) Dalvit, C.; Pevarello, P.; Tatò, M.; Veronesi, M.; Vulpetti, A.; Sundström, M. *J Biomol NMR* **2000**, *18*, 65.
- (138) Antanasijevic, A.; Ramirez, B.; Caffrey, M. *J Biomol NMR* **2014**, *60*, 37.
- (139) Fernandez-Alonso, M. d. C.; Alvaro Berbis, M.; Canales, A.; Arda, A.; Javier Canada, F.; Jimenez-Barbero, J. In *New Applications of NMR in Drug Discovery and Development*; The Royal Society of Chemistry: 2013, p 7.
- (140) Clore, G. M.; Gronenborn, A. M. *Journal of Magnetic Resonance (1969)* **1983**, *53*, 423.
- (141) Dalvit, C.; Flocco, M.; Knapp, S.; Mostardini, M.; Perego, R.; Stockman, B. J.; Veronesi, M.; Varasi, M. *Journal of the American Chemical Society* **2002**, *124*, 7702.
- (142) Dalvit, C.; Fagerness, P. E.; Hadden, D. T. A.; Sarver, R. W.; Stockman, B. J. *Journal of the American Chemical Society* **2003**, *125*, 7696.
- (143) Dalvit, C. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2007**, *51*, 243.
- (144) Aguilar, J. A.; Nilsson, M.; Bodenhausen, G.; Morris, G. A. *Chemical Communications* **2012**, *48*, 811.
- (145) Leung, I. K. H.; Demetriades, M.; Hardy, A. P.; Lejeune, C.; Smart, T. J.; Szöllössi, A.; Kawamura, A.; Schofield, C. J.; Claridge, T. D. W. *Journal of Medicinal Chemistry* **2013**, *56*, 547.
- (146) Hajduk, P. J.; Dinges, J.; Miknis, G. F.; Merlock, M.; Middleton, T.; Kempf, D. J.; Egan, D. A.; Walter, K. A.; Robins, T. S.; Shuker, S. B.; Holzman, T. F.; Fesik, S. W. *Journal of Medicinal Chemistry* **1997**, *40*, 3144.
- (147) van Dongen, M. J. P.; Uppenberg, J.; Svensson, S.; Lundbäck, T.; Åkerud, T.; Wikström, M.; Schultz, J. *Journal of the American Chemical Society* **2002**, *124*, 11874.
- (148) Rademacher, C.; Guiard, J.; Kitov, P. I.; Fiege, B.; Dalton, K. P.; Parra, F.; Bundle, D. R.; Peters, T. *Chemistry – A European Journal* **2011**, *17*, 7442.
- (149) Morris, K. F.; Johnson, C. S. *Journal of the American Chemical Society* **1992**, *114*, 3139.
- (150) Wu, D. H.; Chen, A. D.; Johnson, C. S. *Journal of Magnetic Resonance, Series A* **1995**, *115*, 260.

- (151) Johnson Jr, C. S. *Progress in Nuclear Magnetic Resonance Spectroscopy* **1999**, 34, 203.
- (152) Cameron, K. S.; Fielding, L. *The Journal of Organic Chemistry* **2001**, 66, 6891.
- (153) Wang, H.-B.; Zhao, B.; Zhong, Y.-Q.; Li, K.-C.; Li, Z.-Y.; Wang, Q.; Lu, Y.-J.; Zhang, Z.-N.; He, S.-Q.; Zheng, H.-C.; Wu, S.-X.; Hökfelt, T. G. M.; Bao, L.; Zhang, X. *Proceedings of the National Academy of Sciences of the United States of America* **2010**, 107, 13117.
- (154) Schoffelmeer, A. N.; De Vries, T. J.; Hogenboom, F.; Mulder, A. H. *Journal of Pharmacology and Experimental Therapeutics* **1993**, 267, 205.
- (155) Mignat, C.; Wille, U.; Ziegler, A. *Life Sciences* **1995**, 56, 793.
- (156) Roques, B. P.; Fournié-Zaluski, M.-C.; Wurm, M. *Nature Reviews Drug Discovery* **2012**, 11, 292.
- (157) Wisner, A.; Dufour, E.; Messaoudi, M.; Nejdi, A.; Marcel, A.; Ungeheuer, M.-N.; Rougeot, C. *Proceedings of the National Academy of Sciences* **2006**, 103, 17979.
- (158) Tian, X.-z.; Chen, J.; Xiong, W.; He, T.; Chen, Q. *Peptides* **2009**, 30, 1348.
- (159) Yang, Q.-Z.; Lu, S.-S.; Tian, X.-Z.; Yang, A.-M.; Ge, W.-W.; Chen, Q. *Neuroscience Letters* **2011**, 489, 131.
- (160) Bogeas, A.; Dufour, E.; Bisson, J.-F.; Messaoudi, M.; Rougeot, C. *Biochemistry & Pharmacology* **2013**, 2, 122.
- (161) Fournie-Zaluski, M. C.; Coric, P.; Turcaud, S.; Lucas, E.; Noble, F.; Maldonado, R.; Roques, B. P. *Journal of Medicinal Chemistry* **1992**, 35, 2473.
- (162) Jutkiewicz, E. M. *CNS Drug Reviews* **2007**, 13, 192.
- (163) Noble, F.; Soleilhac, J. M.; Soroca-Lucas, E.; Turcaud, S.; Fournie-Zaluski, M. C.; Roques, B. P. *Journal of Pharmacology and Experimental Therapeutics* **1992**, 261, 181.
- (164) Ruiz, F.; Fournié-Zaluski, M. C.; Roques, B. P.; Maldonado, R. *British Journal of Pharmacology* **1996**, 119, 174.
- (165) Chen, H.; Noble, F.; Coric, P.; Fournie-Zaluski, M.-C.; Roques, B. P. *Proceedings of the National Academy of Sciences* **1998**, 95, 12028.
- (166) Chen, H.; Noble, F.; Mothé, A.; Meudal, H.; Coric, P.; Danascimento, S.; Roques, B. P.; George, P.; Fournié-Zaluski, M.-C. *Journal of Medicinal Chemistry* **2000**, 43, 1398.
- (167) Harvey, J. H.; Long, D. H.; England, P. M.; Whistler, J. L. *ACS Medicinal Chemistry Letters* **2012**, 3, 640.
- (168) Matthews, B. W. *Accounts of Chemical Research* **1988**, 21, 333.
- (169) Roques, B. P.; Noble, F.; Daugé, V.; Fournié-Zaluski, M. C.; Beaumont, A. *Pharmacological Reviews* **1993**, 45, 87.
- (170) Matheson, A.; Noble, S. *Drugs* **2000**, 59, 829.
- (171) Eberlin, M.; Mück, T.; Michel, M. C. *Frontiers in Pharmacology* **2012**, 3, 93.
- (172) Noble, F.; Roques, B. P. *Expert Opinion on Therapeutic Targets* **2007**, 11, 145.
- (173) Cunningham, B. C.; Wells, J. A. *Science* **1989**, 244, 1081.
- (174) Morrison, K. L.; Weiss, G. A. *Current Opinion in Chemical Biology* **2001**, 5, 302.
- (175) Wishart, D.; Bigam, C.; Holm, A.; Hodges, R.; Sykes, B. *J Biomol NMR* **1995**, 5, 67.
- (176) Pinto, M.; Rougeot, C.; Gracia, L.; Rosa, M.; García, A.; Arsequell, G.; Valencia, G.; Centeno, N. B. *ACS Medicinal Chemistry Letters* **2011**, 3, 20.
- (177) Chan, W. C.; White, P. D. *Fmoc Solid Phase Peptide Synthesis: A practical Approach*, 2000.
- (178) Rougeot, C. *PCT/EP2009/054171, April 4, 2009 and WO/2009/124948, October 15, 2009*.
- (179) Mantle, D.; Hardy, M. F.; Lauffart, B.; McDermott, J. R.; Smith, A. I.; Pennington, R. J. *Biochemical Journal* **1983**, 211, 567.
- (180) *MacroModel, version 9.6, Schrödinger, LLC, New York, 2008*.
- (181) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *Journal of the American Chemical Society* **1996**, 118, 11225.

- (182) *Maestro, a powerful all-purposer molecular modeling environment, version 8.5*, Schödingen, LLC, New York, NY 2008.
- (183) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *Journal of the American Chemical Society* **1990**, *112*, 6127.
- (184) Jenck, F.; Moreau, J.-L.; Martin, J. R.; Kilpatrick, G. J.; Reinscheid, R. K.; Monsma, F. J.; Nothacker, H.-P.; Civelli, O. *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 14854.
- (185) Darland, T.; Heinricher, M. M.; Grandy, D. K. *Trends in Neurosciences*, *21*, 215.
- (186) Calo, G.; Guerrini, R.; Rizzi, A.; Salvadori, S.; Regoli, D. *British Journal of Pharmacology* **2000**, *129*, 1261.
- (187) Meunier, J.-C.; Mollereau, C.; Toll, L.; Suaudeau, C.; Moisand, C.; Alvinerie, P.; Butour, J.-L.; Guillemot, J.-C.; Ferrara, P.; Monsarrat, B.; Mazarguil, H.; Vassart, G.; Parmentier, M.; Costentin, J. *Nature* **1995**, *377*, 532.
- (188) Mollereau, C.; Mouledous, L. *Peptides* **2000**, *21*, 907.
- (189) Mollereau, C.; Simons, M. J.; Soularue, P.; Liners, F.; Vassart, G.; Meunier, J. C.; Parmentier, M. *Proceedings of the National Academy of Sciences of the United States of America* **1996**, *93*, 8666.
- (190) Reinscheid, R. K.; Nothacker, H.-P.; Bourson, A.; Ardati, A.; Henningsen, R. A.; Bunzow, J. R.; Grandy, D. K.; Langen, H.; Monsma, F. J.; Civelli, O. *Science* **1995**, *270*, 792.
- (191) Chiou, L. C.; Liao, Y. Y.; Fan, P. C.; Kuo, P. H.; Wang, C. H.; Riemer, C.; Prinssen, E. P. *Current Drug Targets* **2007**, *8*, 117.
- (192) Goeldner, C.; Reiss, D.; Wichmann, J.; Meziane, H.; Kieffer, B. L.; Ouagazzal, A.-M. *The Journal of Neuroscience* **2008**, *28*, 2190.
- (193) Brozovic, A.; Osmak, M. *Cancer Letters* **2007**, *251*, 1.
- (194) Dooley, C. T.; Spaeth, C. G.; Berzetei-Gurske, I. P.; Craymer, K.; Adapa, I. D.; Brandt, S. R.; Houghten, R. A.; Toll, L. *Journal of Pharmacology and Experimental Therapeutics* **1997**, *283*, 735.
- (195) Meng, F.; Taylor, L. P.; Hoversten, M. T.; Ueda, Y.; Ardati, A.; Reinscheid, R. K.; Monsma, F. J.; Watson, S. J.; Civelli, O.; Akil, H. *Journal of Biological Chemistry* **1996**, *271*, 32016.
- (196) Popiolek-Barczyk, K.; Rojewska, E.; Jurga, A. M.; Makuch, W.; Zador, F.; Borsodi, A.; Piotrowska, A.; Przewlocka, B.; Mika, J. *BioMed Research International* **2014**, *762930*.
- (197) Corzana, F.; Fernández-Tejada, A.; Busto, J. H.; Joshi, G.; Davis, A. P.; Jiménez-Barbero, J.; Avenoza, A.; Peregrina, J. M. *ChemBioChem* **2011**, *12*, 110.
- (198) Bosques, C. J.; Tschampel, S. M.; Woods, R. J.; Imperiali, B. *Journal of the American Chemical Society* **2004**, *126*, 8421.
- (199) Norgren, A. S.; Geitmann, M.; Danielson, U. H.; Arvidsson, P. I. *Journal of Molecular Recognition* **2007**, *20*, 132.
- (200) Guerrini, R.; Calo, G.; Rizzi, A.; Bianchi, C.; Lazarus, L. H.; Salvadori, S.; Temussi, P. A.; Regoli, D. *Journal of Medicinal Chemistry* **1997**, *40*, 1789.
- (201) Guerrini, R.; Caló, G.; Bigoni, R.; Rizzi, D.; Rizzi, A.; Zucchini, M.; Varani, K.; Hashiba, E.; Lambert, D. G.; Toth, G.; Borea, P. A.; Salvadori, S.; Regoli, D. *Journal of Medicinal Chemistry* **2001**, *44*, 3956.
- (202) Biondi, B.; Goldin, D.; Giannini, E.; Lattanzi, R.; Negri, L.; Melchiorri, P.; Ciocca, L.; Rocchi, R. *Int J Pept Res Ther* **2006**, *12*, 139.
- (203) Buck, M. *Quarterly Reviews of Biophysics* **1998**, *31*, 297.
- (204) Povey, J. F.; Smales, C. M.; Hassard, S. J.; Howard, M. J. *Journal of Structural Biology* **2007**, *157*, 329.
- (205) Fasman, G. D. In *Circular Dichroism and the Conformational Analysis of Biomolecules*; Springer US: 1996.

- (206) Epand, R. M.; Vogel, H. J. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1999**, *1462*, 11.
- (207) Gao, X.; Wong, T. C. *Biopolymers* **2001**, *58*, 20.
- (208) Marcotte, I.; Auger, M. *Concepts in Magnetic Resonance Part A* **2005**, *24A*, 17.
- (209) Chill, J. H.; Louis, J. M.; Miller, C.; Bax, A. *Protein Science* **2006**, *15*, 684.
- (210) Orsini, M. J.; Nesmelova, I.; Young, H. C.; Hargittai, B.; Beavers, M. P.; Liu, J.; Connolly, P. J.; Middleton, S. A.; Mayo, K. H. *Journal of Biological Chemistry* **2005**, *280*, 8134.
- (211) Yamamoto, T.; Nair, P.; Jacobsen, N. E.; Vagner, J.; Kulkarni, V.; Davis, P.; Ma, S.-w.; Navratilova, E.; Yamamura, H. I.; Vanderah, T. W.; Porreca, F.; Lai, J.; Hruby, V. J. *Journal of Medicinal Chemistry* **2009**, *52*, 5164.
- (212) Palian, M. M.; Boguslavsky, V. I.; O'Brien, D. F.; Polt, R. *Journal of the American Chemical Society* **2003**, *125*, 5823.
- (213) Yamamoto, T.; Nair, P.; Jacobsen, N. E.; Davis, P.; Ma, S.-w.; Navratilova, E.; Moye, S.; Lai, J.; Yamamura, H. I.; Vanderah, T. W.; Porreca, F.; Hruby, V. J. *Journal of Medicinal Chemistry* **2008**, *51*, 6334.
- (214) Mukerjee, P.; Mysels, K. J. *Journal of Pharmaceutical Sciences* **1972**, *61*, 319.
- (215) Israelachvili, J. N. *Boston: Academic Press* **1992**.
- (216) Wüthrich, K. *NMR of Proteins and Nucleic Acids*. Wiley, New York. **1986**.
- (217) Corzana, F.; Busto, J. H.; Jiménez-Osés, G.; García de Luis, M.; Asensio, J. L.; Jiménez-Barbero, J.; Peregrina, J. M.; Avenoza, A. *Journal of the American Chemical Society* **2007**, *129*, 9458.
- (218) Fernández-Tejada, A.; Corzana, F.; Busto, J. H.; Jiménez-Osés, G.; Jiménez-Barbero, J.; Avenoza, A.; Peregrina, J. M. *Chemistry – A European Journal* **2009**, *15*, 7297.
- (219) Wishart, D.; Bigam, C.; Yao, J.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J.; Sykes, B. *J Biomol NMR* **1995**, *6*, 135.
- (220) Pratt, M. R.; Bertozzi, C. R. *Chemical Society Reviews* **2005**, *34*, 58.
- (221) Kogelberg, H.; Solis, D.; Jimenez-Barbero, J. *Current Opinion in Structural Biology* **2003**, *13*, 646.
- (222) Harrison, R. S.; Ruiz-Gómez, G.; Hill, T. A.; Chow, S. Y.; Shepherd, N. E.; Lohman, R.-J.; Abbenante, G.; Hoang, H. N.; Fairlie, D. P. *Journal of Medicinal Chemistry* **2010**, *53*, 8400.
- (223) Winterfeld, Gottfried A.; Khodair, Ahmed I.; Schmidt, Richard R. *European Journal of Organic Chemistry* **2003**, *2003*, 1009.
- (224) Mayato, C.; Dorta, R. L.; Vázquez, J. T. *Tetrahedron Letters* **2008**, *49*, 1396.
- (225) Arsequell, G.; Krippner, L.; Dwek, R. A.; Wong, S. Y. C. *Journal of the Chemical Society, Chemical Communications* **1994**, *0*, 2383.
- (226) Günther, P. *Methods in Molecular Biology* **2004**, *278*, 353
- (227) Koradi, R.; Billeter, M.; Wüthrich, K. *Journal of Molecular Graphics* **1996**, *14*, 51.
- (228) Gabius, H.-J. *International Journal of Biochemistry* **1994**, *26*, 469.
- (229) Sharon, N.; Lis, H. *Glycobiology* **2004**, *14*, 53R.
- (230) Sharon, N. *Journal of Biological Chemistry* **2007**, *282*, 2753.
- (231) Kronis, K. A.; Carver, J. P. *Biochemistry* **1982**, *21*, 3050.
- (232) Aub, J. C.; Sanford, B. H.; Wang, L. H. *Proceedings of the National Academy of Sciences of the United States of America* **1965**, *54*, 400.
- (233) Burger, M. M.; Goldberg, A. R. *Proceedings of the National Academy of Sciences of the United States of America* **1967**, *57*, 359.
- (234) Jacobs, S.; Shechter, Y.; Bissell, K.; Cuatrecasas, P. *Biochemical and Biophysical Research Communications* **1977**, *77*, 981.
- (235) Zhao, J.; Simeone, D. M.; Heidt, D.; Anderson, M. A.; Lubman, D. M. *Journal of Proteome Research* **2006**, *5*, 1792.
- (236) Lycknert, K.; Edblad, M.; Imberty, A.; Widmalm, G. *Biochemistry* **2004**, *43*, 9647.

- (237) Wright, C. S. *Journal of Molecular Biology* **1980**, 141, 267.
- (238) Wright, C. S. *Journal of Molecular Biology* **1989**, 209, 475.
- (239) Schwefel, D.; Maierhofer, C.; Beck, J. G.; Seeberger, S.; Diederichs, K.; Möller, H. M.; Welte, W.; Wittmann, V. *Journal of the American Chemical Society* **2010**, 132, 8704.
- (240) Wright, C. S.; Kellogg, G. E. *Protein Science : A Publication of the Protein Society* **1996**, 5, 1466.
- (241) Wright, C. S. *Journal of Molecular Biology* **1990**, 215, 635.
- (242) Wright, C. S.; Jaeger, J. *Journal of Molecular Biology* **1993**, 232, 620.
- (243) Harata, K.; Nagahora, H.; Jigami, Y. *Acta Crystallographica Section D* **1995**, 51, 1013.
- (244) Muraki, M.; Ishimura, M.; Harata, K. *Biochimica et Biophysica Acta* **2002**, 1569, 10.
- (245) Espinosa, J. F.; Asensio, J. L.; García, J. L.; Laynez, J.; Bruix, M.; Wright, C.; Siebert, H. C.; Gabius, H. J.; Cañada, F. J.; Jiménez-Barbero, J. *European Journal of Biochemistry* **2000**, 267, 3965.
- (246) Privat, J.-P.; Delmotte, F.; Mialonier, G.; Bouchard, P.; Monsigny, M. *European Journal of Biochemistry* **1974**, 47, 5.
- (247) Garnett, J. A.; Liu, Y.; Leon, E.; Allman, S. A.; Friedrich, N.; Saouros, S.; Curry, S.; Soldati-Favre, D.; Davis, B. G.; Feizi, T.; Matthews, S. *Protein Science : A Publication of the Protein Society* **2009**, 18, 1935.
- (248) Lycknert, K.; Edblad, M.; Imberty, A.; Widmalm, G. *Biochemistry* **2004**, 43, 9647.
- (249) Lotan, R.; Sharon, N.; Mirelman, D. *European Journal of Biochemistry* **1975**, 55, 257.
- (250) Mrkoci, K.; Kelm, S.; Crocker, P.; Schauer, R.; Berger, E. *Glycoconjugate J* **1996**, 13, 567.
- (251) Braun, P.; Nägele, B.; Wittmann, V.; Drescher, M. *Angewandte Chemie International Edition* **2011**, 50, 8428.
- (252) Wittmann, V. *Current Opinion in Chemical Biology* **2013**, 17, 982.
- (253) Wittmann, V.; Pieters, R. J. *Chemical Society Reviews* **2013**, 42, 4492.
- (254) Ghirardello, M.; Öberg, K.; Staderini, S.; Renaudet, O.; Berthet, N.; Dumy, P.; Hed, Y.; Marra, A.; Malkoch, M.; Dondoni, A. *Journal of Polymer Science Part A: Polymer Chemistry* **2014**, 52, 2422.
- (255) Gouverneur, V.; Muller, K. In *Molecular medicine and medicinal chemistry*; London: Imperial College Press, 2012.
- (256) Purser, S.; Moore, P. R.; Swallow, S.; Gouverneur, V. *Chemical Society Reviews* **2008**, 37, 320.
- (257) Müller, K.; Faeh, C.; Diederich, F. *Science* **2007**, 317, 1881.
- (258) Wang, J.; Sánchez-Roselló, M.; Aceña, J. L.; del Pozo, C.; Sorochnikov, A. E.; Fustero, S.; Soloshonok, V. A.; Liu, H. *Chemical Reviews* **2014**, 114, 2432.
- (259) Diercks, T.; Ribeiro, J. P.; Cañada, F. J.; André, S.; Jiménez-Barbero, J.; Gabius, H.-J. *Chemistry – A European Journal* **2009**, 15, 5666.
- (260) Asensio, J. L.; Cañada, F. J.; Bruix, M.; González, C.; Khiar, N.; Rodríguez-Romero, A.; Jiménez-Barbero, J. *Glycobiology* **1998**, 8, 569.
- (261) Jiménez-Barbero, J.; Javier Cañada, F.; Asensio, J. L.; Aboitiz, N.; Vidal, P.; Canales, A.; Groves, P.; Gabius, H. J.; Siebert, H. C. C. C. B. *Advances in Carbohydrate Chemistry and Biochemistry* **2006**, 60.
- (262) Beckmann, H. S. G.; Möller, H. M.; Wittmann, V. *Beilstein Journal of Organic Chemistry* **2012**, 8, 819.
- (263) Asensio, J. L.; Cañada, F. J.; Bruix, M.; González, C.; Khiar, N.; Rodríguez-Romero, A.; Jiménez-Barbero, J. *Glycobiology* **1998**, 8, 569.

- (264) Asensio, J. L.; Siebert, H. C.; von Der Lieth, C. W.; Laynez, J.; Bruix, M.; Soedjanaamadja, U. M.; Beintema, J. J.; Cañada, F. J.; Gabius, H. J.; Jiménez-Barbero, J. *Proteins* **2000**, *40*, 218.
- (265) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *Journal of the American Chemical Society* **1995**, *117*, 5179.
- (266) Dubowchik, G. M.; Vrudhula, V. M.; Dasgupta, B.; Ditta, J.; Chen, T.; Sheriff, S.; Sipman, K.; Witmer, M.; Tredup, J.; Vyas, D. M.; Verdoorn, T. A.; Bollini, S.; Vinitzky, A. *Organic Letters* **2001**, *3*, 3987.
- (267) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; Gaussian, Inc.: Wallingford, CT, USA, 2009.
- (268) Boys, S. F.; Bernardi, F. *Molecular Physics* **1970**, *19*, 553.
- (269) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *Journal of Chemical Theory and Computation* **2006**, *2*, 364.
- (270) Zhao, Y.; Truhlar, D. G. *Accounts of Chemical Research* **2008**, *41*, 157.
- (271) Paizs, B.; Suhai, S. *Journal of Computational Chemistry* **1998**, *19*, 575.
- (272) Reyes-López, C. A.; Hernández-Santoyo, A.; Pedraza-Escalona, M.; Mendoza, G.; Hernández-Arana, A.; Rodríguez-Romero, A. *Biochemical and Biophysical Research Communications* **2004**, *314*, 123.
- (273) Aboitiz, N.; Vila-Perelló, M.; Groves, P.; Asensio, J. L.; Andreu, D.; Cañada, F. J.; Jiménez-Barbero, J. *Chembiochem* **2004**, *5*, 1245.
- (274) Krishnan, S.; Hall, B. G.; Sinnott, M. L. *Biochemical Journal* **1995**, *312*, 971.
- (275) Juers, D. H.; Matthews, B. W.; Huber, R. E. *Protein Science : A Publication of the Protein Society* **2012**, *21*, 1792.
- (276) Banner, D. W.; Bloomer, A. C.; Petsko, G. A.; Phillips, D. C.; Wilson, I. A. *Biochemical and Biophysical Research Communications* **1976**, *72*, 146.
- (277) Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P. M.; Henrissat, B. *Nucleic Acids Research* **2014**, *42*, D490.
- (278) Juers, D. H.; Heightman, T. D.; Vasella, A.; McCarter, J. D.; Mackenzie, L.; Withers, S. G.; Matthews, B. W. *Biochemistry* **2001**, *40*, 14781.
- (279) Bader, D. E.; Ring, M.; Huber, R. E. *Biochemical and Biophysical Research Communications* **1988**, *153*, 301.
- (280) Ring, M.; Bader, D. E.; Huber, R. E. *Biochemical and Biophysical Research Communications* **1988**, *152*, 1050.
- (281) Spiwok, V.; Lipovová, P.; Skálová, T.; Buchtelová, E.; Hasek, J.; Králová, B. *Carbohydrate Research* **2004**, *339*, 2275.
- (282) Wheatley, R. W.; Lo, S.; Jancewicz, L. J.; Dugdale, M. L.; Huber, R. E. *Journal of Biological Chemistry* **2013**.
- (283) Hadfield, A. T.; Harvey, D. J.; Archer, D. B.; MacKenzie, D. A.; Jeenes, D. J.; Radford, S. E.; Lowe, G.; Dobson, C. M.; Johnson, L. N. *Journal of Molecular Biology* **1994**, *243*, 856.

- (284) Cagnoni, A. J.; Uhrig, M. L.; Varela, O. *Bioorganic & Medicinal Chemistry* **2009**, *17*, 6203.
- (285) Navarro-Vázquez, A.; Cobas, J. C.; Sardina, F. J.; Casanueva, J.; Díez, E. *Journal of Chemical Information and Computer Sciences* **2004**, *44*, 1680.
- (286) Cantor, C. R.; Schimmel, P. R. *Biophysical chemistry: Part II 'Techniques for the study of biological structure and function'* **1980**, WH Freeman and Company: San Francisco, CA., p. 461.
- (287) Vandenbussche, S.; Díaz, D.; Fernández-Alonso, M. C.; Pan, W.; Vincent, S. P.; Cuevas, G.; Cañada, F. J.; Jiménez-Barbero, J.; Bartik, K. *Chemistry – A European Journal* **2008**, *14*, 7570.
- (288) Biarnés, X.; Ardèvol, A.; Planas, A.; Rovira, C.; Laio, A.; Parrinello, M. *Journal of the American Chemical Society* **2007**, *129*, 10686.
- (289) Lammerts van Bueren, A.; Ardèvol, A.; Fayers-Kerr, J.; Luo, B.; Zhang, Y.; Sollogoub, M.; Blériot, Y.; Rovira, C.; Davies, G. J. *Journal of the American Chemical Society* **2010**, *132*, 1804.
- (290) Gloster, T. M.; Davies, G. J. *Organic and Biomolecular Chemistry* **2010**, *8*, 305.
- (291) Uhrig, M. L.; Manzano, V. E.; Varela, O. *European Journal of Organic Chemistry* **2006**, *2006*, 162.
- (292) Strino, F.; Lii, J.-H.; Gabius, H.-J.; Nyholm, P.-G. *J Comput Aided Mol Des* **2009**, *23*, 845.
- (293) *Glide, version 5.5, Schrödinger, Inc., New York, NY, 2009.*
- (294) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *Journal of Computational Chemistry* **2005**, *26*, 1668.
- (295) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; Gonzalez-Outeirino, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. *Journal of Computational Chemistry* **2008**, *29*, 622.
- (296) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *Journal of Computational Physics* **1977**, *23*, 327.
- (297) York, D. M.; Darden, T. A.; Pedersen, L. G. *The effect of long-range electrostatic interactions in simulations of macromolecular crystals: A comparison of the Ewald and truncated list methods*; AIP, 1993; Vol. 99.
- (298) Humphrey, W.; Dalke, A.; Schulten, K. *Journal of Molecular Graphics* **1996**, *14*, 33.
- (299) Tian, X. Z.; Chen, J.; Xiong, W.; He, T.; Chen, Q. *Peptides* **2009**, *30*, 1348.
- (300) Lambert, D. G. *Nature Reviews Drug Discovery* **2008**, *7*, 694.
- (301) Jacobson, R. H.; Zhang, X. J.; DuBose, R. F.; Matthews, B. W. *Nature* **1994**, *369*, 761.